Engaging the People to Look Beyond the Surface of Online Information Visualizations of Open Data

Jeremy Boy

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2015-ENST-0025

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Sciences de l'Information et de la Communication »

présentée et soutenue publiquement par

Jeremy BOY

le 18 mai 2015

Engager les Citoyens à Aller au-delà des Simples Représentations de Données Ouvertes

Co-directrice de thèse : **Françoise DETIENNE**
Co-directeur de thèse : **Jean-Daniel FEKETE**

Jury

Mme Wendy MACKAY, Directrice de recherche, In-Situ, Inria Saclay	Directrice du jury
M. Stéphane CONVERSY, Maître de conférence HDR, LII, ENAC	Rapporteur
M. Jo WOOD, Professor, giCenter, City University London	Rapporteur
M. James EAGAN, Maître de conférence, VIA-INFRES, TELECOM ParisTech	Examinateur
Mme Françoise DETIENNE, Directrice de recherche, LTCI-CNRS, TELECOM ParisTech	Co-directrice
M. Jean-Daniel FEKETE, Directeur de recherche, Aviz, Inria Saclay	Co-directeur
M. Remy BOURGANEL, Enseignant, Sociable Media, EnsadLab	Invité

T H È S E

TELECOM ParisTech
école de l'Institut Mines-Télécom - membre de ParisTech
46 rue Barrault 75013 Paris - (+33) 1 45 81 77 77 - www.telecom-paristech.fr

Abstract

In this dissertation, I explore four initial challenges an online user may encounter when confronted with an information visualization website. These challenges occur in the early stages of data-exploration, and may prevent the user—who *a priori* has no specific training—from engaging in the process. As engaging crowds of both experts and non-experts—be it in visualization or in the data-topic—in social data-analysis has well recognized applications and benefits, it is necessary to consider how well these people understand and use visualizations in order to improve them. The main research question I address is: how might the different challenges people may encounter limit their engagement in efficient explorations of data, and how might these limitations be remedied? To answer this, I first provide some context to emphasize the importance of finding ways to enable people to best take advantage of the possibilities provided by online visualisations—especially in the case of visualizations of open data. Next, I define the four initial challenges as sub-costs of van Wijk's perception and exploration costs associated with using a visualization technique; these sub-costs are: 1) a literacy cost, 2) a context-interpretation cost, 3) a perceived interactivity cost, and 4) an initial incentive for exploration cost. I then set four lower-level research questions (one for each cost), which I address in the main chapters of this dissertation. For each, I propose either a way to assess or a method to help overcome the sub-cost. I also investigate whether popular techniques recommended for making visualizations engaging outside of purely analytical contexts can lead online users to explore data. The results of this work show that each cost is indeed an important barrier to the engagement of online users in data-explorations; and they encourage the pursuit of research on the different design methods I develop to help people overcome these challenges.

Résumé

Dans ce manuscrit, j'étudie quatre obstacles potentiels à l'engagement d'un internaute avec une interface de visualisation d'informations interactive. Si cet utilisateur n'a pas de connaissances particulières et n'a pas reçu de formation initiale sur l'utilisation d'outils de visualisations, il peut avoir du mal à se plonger dans l'exploration efficace d'un jeu de données — même s'il est intéressé par le sujet. Il est donc nécessaire de questionner la capacité qu'a cet internaute à comprendre différents types de visualisations afin d'en améliorer le design. Ceci peut favoriser l'analyse sociale de données et, dans le cas des données ouvertes, enrichir le débat public. Ma question de recherche principale est : comment les différents obstacles auquel un internaute peut être confronté sont-ils susceptibles de limiter son engagement dans l'exploration personnelle et efficace de données et comment remédier à ces limitations ? Pour y répondre, je commence par contextualiser l'importance du rôle que peut jouer la visualisation sur le web, surtout dans le cadre des données ouvertes. Ensuite, je définis quatre obstacles à l'engagement en termes de sous-coûts de la perception et de l'exploration liés à l'utilisation d'une interface de visualisation en me référant au modèle proposé par van Wijk. Ces sous-coûts sont : 1) un coût de littératie, 2) un coût d'interprétation du contexte, 3) un coût de perception d'interactivité et 4) un coût de motivation initiale à explorer des données. Ceux-ci me permettent alors de décomposer et d'opérationnaliser ma question de recherche en quatre sous-questions spécifiques (relatives à chaque coût) auxquelles je réponds dans les chapitres principaux de ce manuscrit. Pour chacun, j'adopte soit une approche expérimentale pour mesurer le coût en question, soit une approche design pour aider les internautes à le surmonter. J'évalue aussi l'effet de certains éléments de design de visualisation reconnus pour leurs qualités communicationnelles sur l'engagement des internautes à explorer des données. Les résultats de ce travail montrent que chaque coût peut effectivement être un obstacle important à l'engagement d'un internaute ; ils encouragent à poursuivre cette recherche à partir des différentes méthodes de design que je propose.

Acknowledgements

First and foremost, I want to thank Jean-Daniel Fekete and Françoise Detienne for having advised me during these past three years, and for believing in my potential. Seeing other people around me start a PhD, and seeing how difficult a process it actually is, makes me realize how lucky I was when I first met these great people, and how much trust they invested in me for things to get set up so quickly. I also thank Jean-François Depelsenaire who helped me get started on this adventure, as well as Remy Bourganel who supported me during these three years—I apologize that things ended the way they did. Finally, I thank Google and Inria for supporting my work financially.

My thoughts then go out to the other great people I have met and had the chance to work with during my PhD. First, to the people in the Aviz team at Inria who are all brilliant researchers and good people to be around. Thanks to Mathieu Le Goc, Samuel Huron, Charles Perin, Andre Spritzer and Romain Vuillemot who I consider good friends—always looking forward to the next conference so we can meet up again; to Pierre Dragicevic, Petra Isenberg, and Tobias Isenberg for sharing their knowledge and wisdom with me; to Benjamin Bach, Yvonne Jansen, Pascal Goffin, Wesley Willett, Lora Oehlberg, Anthi Dimara, Nadia Boukhelifa, and Nicolat Heulot who have also taught me a lot in different ways; and finally, to Romain Di Vozzo for listening to me blabber on about art and design—sorry for those I missed out. Second, to the people in the Sociable Media group at EnsadLab who are great designers, interesting people, and close friends. Thanks to Pauline Gourlet, Max Mollon, Louis Eveillard, Ferdinand Dervieux, Sarah Garcin, and Emeline Brulé for supporting me until the end, for great culinary experiences, and probably for one too many drinks. Third, to the people in the SES department at Telecom ParisTech with whom I have always had interesting discussions that have challenged my perspectives. Thanks to Annie Gentes and Michael Baker. Finally fourth, to my different collaborators who have greatly contributed to my evolution as a researcher. Thanks to Ronald A. Rensink, Enrico Bertini, Sung Hee Kim, Sukwon

Lee, Ji Soo Yi, and Niklas Elmqvist for sharing great ideas with me.

On a more personal note, I want to thank my mum and dad who have always been supportive of my choices, and who, I am now glad to announce, I better understand thanks to my initial journey as a researcher—at least I now have clearer idea of what they do for a living. Also, thanks to my sister for challenging my views on the world; to Ludovic Bruggeman for being a great friend and for always helping me disconnect with the work overload—probably also related to the way too many drinks we have had together—as well as to Benjamin Roussel and Émilie Robin. Finally, thanks to my friends in Toulouse who have also been of great help for seeing things differently—through rollerblading, music, *etc.*

To finish, I deeply thank Mathilde Camagne for her kind and loving support over the years, and I am truly sorry and affected by the way things turned out. I hope we will remain close…

Contents

Figures

Tables

Résumé Étendu

Dans ce manuscrit, j'étudie quatre obstacles potentiels à l'engagement d'un internaute avec une interface de *visualisation d'informations* sur le web. Je me concentre sur des graphiques interactifs qui mettent en forme des *données ouvertes*, c'est-à-dire de données rendues accessibles sur Internet par des gouvernements et autres organisations publiques. Ces données peuvent enrichir le débat public et sont aujourd'hui envisagées comme un nouvel outil de démocratie. Toutefois, réussir à les comprendre et à en extraire du sens n'est pas une tâche triviale. C'est justement l'objectif de la visualisation de données et c'est la raison pour laquelle elle est aujourd'hui massivement utilisée, notamment sur Internet. Or la plupart de ces visualisations sont développées à des fins médiatiques, à partir de données préalablement traitées par des analystes ou des journalistes, ce qui limite le potentiel individuel et social de compréhension et d'exploration des données ouvertes ; et par conséquent, les divergences de perspectives possibles dans le débat public. J'estime qu'il est donc important de trouver des moyens d'engager les citoyens à utiliser des visualisations qui leur permettent d'explorer les données ouvertes *par* et *pour* eux-mêmes afin d'éveiller les consciences sur les faits du monde qui nous entoure. Il s'agit alors de trouver des moyens d'aider ces utilisateurs à dépasser les obstacles auxquels ils peuvent être confrontés.

Ce manuscrit est destiné à deux audiences. Tout d'abord, il est destiné aux *chercheurs* en visualisation d'informations (infovis) puisqu'il présente une série d'expérimentations et d'études menées soit pour identifier et comprendre les obstacles évoqués ci-dessus (Chapter 3), soit pour évaluer l'impact de certains choix de design dans l'engagement des internautes à explorer les données ouvertes (Chapter 5 et Chapter 6). Ensuite, il est destiné aux *designers* puisqu'il introduit différents modèles conceptuels pour le design de visualisations (Chapter 4) et il montre comment l'étude de ces obstacles peut aider le processus de design (Chapter 3 et Chapter 7).

Enfin, ce manuscrit est rédigé en anglais, raison pour laquelle je fournis ici ce résumé étendu en français. Celui-ci est globalement structuré de la même manière que le manuscrit original : chaque section correspond à un chapitre. Dans l'introduction, je présente le contexte général : je décris ce que sont les données ouvertes, ainsi que les enjeux associés à leur impact potentiel ; je fais un court descriptif de ce qu'est la visualisation d'informations ; j'établis ma question de recherche principale ; et j'identifie les quatre obstacles (que je qualifie en termes de *coûts* liés à l'utilisation d'une visualisation). Ceux-ci me permettent de décomposer et d'opérationnaliser ma question de recherche en sous-questions spécifiques à chacun(1). Ensuite, pour les quatre chapitres principaux du manuscrit — qui correspondent respectivement aux différents coûts identifiés — je commence par donner un élément de contexte (par exemple un extrait d'article publié dans la presse ou sur un blog) ; puis j'introduis la sous-question de recherche spécifique au coût adressé ; je liste les contributions scientifiques ; et je discute les résultats. Enfin, dans la conclusion j'expose certaines limitations du travail effectué ; je réponds à ma question de recherche principale ; et je propose plusieurs pistes pour continuer cette recherche.

1 Il est à noter que je ne résume pas mon état de l'art (présenté dans le Chapter 2) ici.

I Introduction

Dans un article posté en novembre 2009, Steve Mollman évoquait déjà la quantité sans cesse croissante de données disponibles sur Internet et des nouveaux enjeux associés à leur compréhension [206]. Selon lui, la « bonne nouvelle » était que de plus en plus de gouvernements et autres organisations publiques ouvraient leur données sur Internet; la « mauvaise nouvelle » était que ces données étaient plutôt austères. Ceci entraînait un intérêt grandissant pour la visualisation de données qui, selon Mollman, « peut transformer des statistiques ennuyeuses en représentations graphiques attrayantes expliquant le monde qui nous entoure. »[2]

Bien que la visualisation soit reconnue comme étant un moyen efficace pour rendre des données plus intelligibles — et qu'elle soit devenue un média inévitable pour les communiquer — la visualisation d'informations permet aussi d'explorer des données grâce à diverses techniques d'interaction, ce qui permet d'en extraire des informations nouvelles. Or cet aspect semble souvent négligé dans les visualisations adressées au grand public.

Cette section résume le Chapter 1. Elle présente d'abord le contexte général du manuscrit en décrivant le concept de données ouvertes et en définissant ce qu'est la visualisation d'informations. Ensuite, elle introduit la question de recherche principale de cette thèse et mon hypothèse générale. Enfin, elle décrit les quatre obstacles à l'engagement d'un internaute avec une interface de visualisation.

I.I Contexte

Les données ouvertes sont à la mode. On compte aujourd'hui sept fois plus de portails gouvernementaux donnant accès à des données qu'en 2010 (56 en 2014 contre seulement 8 en 2010) et presque quatre fois plus de jeux de données à télécharger (431 585 en 2014 contre 138 242 en 2010). Ces don-

2 Traductions personnelles de l'anglais.

nées sont caractérisées par le fait qu'elles sont *universellement accessibles, réutilisables et re-distribuables, sans aucune restriction légale, technologique ou sociale* ; elles sont les « briques de la connaissance ouverte » [103] et traitent principalement de sujets culturels, politiques, scientifiques, financiers, statistiques, météorologiques, environnementaux ou relatifs aux transports.

L'intérêt des gouvernements et autres organisations publiques est de promouvoir la transparence et la participation citoyenne, ainsi que de favoriser l'innovation sociale, économique et scientifique ; mais, malgré la noblesse de ces intentions, il est difficile de savoir à quel point ces données ont un impact réel à l'échelle citoyenne puisqu'elles sont complexes à interpréter et à analyser. Si l'objectif est vraiment d'éveiller les consciences, les citoyens doivent pouvoir comprendre et utiliser ces données par et pour eux-même. Malheureusement en l'état, ceci requiert un degré d'expertise que peu ont. Ainsi, bien que ces données respectent les conventions de l'open framework [56], dans le sens où elle sont distribuées dans des formats lisibles et modifiables par ordinateur sans restriction technologique, elles restent ‹ fermées › pour une majorité de citoyens puisqu'elles sont difficilement ‹ lisibles › pour des êtres humains.

Or plus de vingt-cinq ans de recherche en visualisation d'informations (infovis) ont montré que la visualisation est un moyen efficace pour rendre de telles données plus intelligibles ; et aujourd'hui elle est « déployée contre cet assaut d'informations pour nous aider à comprendre efficacement le déluge numérique »[3] [147].

La visualisation d'informations est définie par *« l'utilisation de représentations visuelles numériques et interactives de données abstraites pour amplifier la cognition »* [131], où les données *abstraites* sont des données qui n'ont pas de représentation canonique. Comme nombre d'autres types de représentations visuelles, la visualisation est utile pour comprendre et communiquer des informations puisqu'elle permet d'accroître le nombre d'objets qu'un individu peut considérer simultanément et elle en favorise le traitement cognitif. Ceci peut alors mener à de nouvelles déductions, inférences ou découvertes [261],

3 Traduction personnelle de l'anglais.

et ainsi aider à donner du sens aux objets représentés et à prendre des décisions éclairées. Toutefois, la visualisation diffère d'autres types de représentations plus symboliques (comme le texte/la typographie) puisqu'elle repose principalement sur des propriétés visuelles *préattentives* qui permettent une lecture très rapide (entre 200 et 250 millisecondes) et précise [81]. De plus, la visualisation d'informations est interactive, ce qui veut dire que les données et leur représentation visuelle peuvent être manipulées pour mieux convenir aux besoins analytiques ou de communication de l'utilisateur.

Mais l'utilisation d'une visualisation d'informations peut aussi être coûteuse. van Wijk [263] identifie les quatre coûts suivants[4], associés à la création et à l'utilisation d'une technique de visualisation :

* Ci — **le coût initial de développement :** chaque technique de visualisation doit être implémentée et un équipement informatique spécifique peut être requis pour ce faire ;
* Cu — **le coût initial par utilisateur :** l'utilisateur doit choisir et acquérir une technique de visualisation et il doit apprendre à s'en servir pour l'ajuster à ses besoins ;
* Cs — **le coût initial par session :** des données doivent être récupérées, traitées et intégrées à la technique de visualisation ; et
* Ce — **les coûts de la perception et de l'exploration :** l'utilisateur doit comprendre la représentation visuelle et doit apprendre à interagir avec pour explorer les données.

Ces coûts sont intégrés dans un *modèle économique* établi par van Wijk pour déterminer si une technique de visualisation vaut la peine d'être utilisée. Le modèle peut se résumer à une différence entre le retour sur investissement et les coûts de cette technique. Si la différence est positive, c'est-à-dire si le retour sur investissement est supérieur aux coûts, alors la technique est intéressante ; sinon elle ne l'est pas.

4 Le nom de chaque coût est ici une traduction personnelle de l'anglais.

Ainsi, la visualisation d'informations n'est parfois pas aussi efficace et engageante que Mollman le laisserait entendre [206]. Dans ce manuscrit, j'identifie quatre raisons potentielles à cela. Premièrement, au-delà des difficultés techniques liées à l'informatique — c'est-à-dire Ci, Cu et Cs — il se peut qu'un utilisateur ne comprenne pas le *langage visuel* utilisé — c'est-à-dire Ce —soit parce qu'il ne sait pas interpréter une visualisation comme une représentation de données, soit parce qu'il ne comprend pas comment en extraire de l'information. Deuxièmement, ce langage est globalement très abstrait et n'est pas toujours très attirant. Les diagrammes habituellement utilisés en entreprise, comme les *diagrammes à barres*, *les diagrammes linéaires* et *les camemberts* — plus généralement tous les diagrammes produits en quelques clicks dans Microsoft Excel qui sont malheureusement très communs — sont assez ternes et peuvent ne pas attiser la curiosité d'un utilisateur. Troisièmement, la visualisation d'informations permet aussi d'interagir avec les données pour les explorer, ce qui est particulièrement utile lorsque celles-ci sont nombreuses et complexes. Ceci nécessite une sensibilisation au potentiel interactif des visualisations, surtout sur Internet. Enfin quatrièmement, comme il se peut qu'un internaute manque de connaissances sur le sujet traité par jeu de données ouvertes, il est possible qu'il ait du mal à formuler des questions initiales suffisamment intéressantes pour lui donner envie d'explorer les données. Ceci pourrait alors l'empêcher d'entamer un processus d'analyse et de compréhension des données par et pour lui-même.

Il faut donc avant tout identifier l'audience à laquelle sont adressées les visualisations de données ouvertes, ainsi que comprendre le contexte dans lequel elles sont déployées.

I.II Quatre sous-coûts de la perception et de l'exploration

La plupart des internautes ne sont ni designer ni développeur informatique. De la même manière, peu sont experts en analyse de données ou en statistiques. Je pense donc qu'il est nécessaire de s'assurer que ces personnes comprennent l'intérêt de la visualisation et sachent utiliser des représentations visuelles pour interpréter et analyser des données. Pour ce faire, je décompose les coûts de la perception et de l'exploration identifiés par van Wijk (Ce) afin de mieux les comprendre et pour pouvoir établir des stratégies de design adéquates. Ma question de recherche principale est donc la suivante :

Comment les coûts de la perception et de l'exploration peuvent-ils limiter l'engagement d'un internaute dans l'exploration personnelle et efficace de données et comment remédier à ces limitations ?

Afin d'opérationnaliser cette question, je propose tout d'abord de mieux définir les audiences auxquelles les visualisations d'informations de données ouvertes sont adressées. Ensuite, je décris les difficultés liées à leur déploiement sur Internet. Enfin, j'identifie quatre *sous-coûts* de Ce (les quatre obstacles mentionnés plus haut) à l'aide d'une analogie avec la théorie du *butinage d'informations*(5).

Je définis les audiences de la visualisation de données ouvertes comme des audiences *occasionnelles*, par opposition aux audiences *expertes.* Je les distingue de ces dernières de la manière suivante : les audiences occasionnelles sont 1) des individus qui sont généralement confrontés à la visualisation en tant que média sur Internet (et non en tant qu'outil de travail), ce qui laisse supposer qu'ils les scruteront certainement à la va-vite, si tant est qu'ils les regardent ; et 2) des individus qui n'ont pas nécessairement de connaissances particulières sur le sujet ou le domaine traité par les données, ni sur les systèmes de visualisation d'informations.

Concernant 1), il est bien connu que l'attention d'un internaute est généralement très limitée ; Nielsen estime que le temps moyen passé sur une page web est inférieur à une minute [28]. Toutefois, Liu *et al.* [192] ont montré que la navigation sur internet est sujette à un *effet de vieillissement négatif*(6), c'est-à-dire que les internautes ont tendance à scanner rapide-

5 Traduction personnelle de l'anglais *information foraging.*

6 Traduction personnelle de l'anglais *negative aging.*

ment une page web avant de décider de ‹ creuser › pour y trouver de l'information. Durant cette phase, la probabilité qu'ils quittent la page est très forte, mais s'ils la dépassent, cette probabilité diminue et les internautes passeront plus de temps à lire et à explorer le contenu de la page.

Concernant 2), un des problèmes spécifiques de la visualisation sur Internet est qu'elle permet généralement des interactions bien plus riches — et donc complexes — que les traditionnels scrolls et clicks sur des hyperliens des pages ‹ classiques ›. Il faut donc que l'internaute soit au courant de ce potentiel interactif, ainsi que de sa richesse. S'il ne comprend pas l'interactivité, il se peut qu'il ne perçoive pas le média comme étant intéressant et il en préférera sûrement un autre (comme du texte par exemple), même si le sujet des données est susceptible de l'intéresser.

Ainsi, j'insiste sur l'importance de considérer Ce lors de la conception de visualisations pour des audiences occasionnelles. J'estime que ces coûts sont une des raisons principales pour lesquelles un internaute pourrait négliger un site de visualisations; celui-ci doit donc non-seulement être *attirant* visuellement, il doit aussi *engager* l'utilisateur à interagir — afin que ce dernier puisse aller au-delà de la surface de la représentation visuelle.

La théorie du butinage d'informations [222] suggère que des individus à la recherche d'informations sont continuellement en train de naviguer entre des *lopins d'informations*[7], guidés par *l'odeur de l'information*[8]. Ces lopins sont des régions dans lesquelles l'information est compilée (par exemple des sites web) et l'odeur est l'estimation faite par l'internaute du potentiel informatif de chaque lopin. Cet *odorat* se développe par la pratique et l'expérience, mais l'odeur peut être renforcée par le design. Selon Nielsen, l'idée principale derrière la théorie du butinage d'informations est l'analyse des coûts et des bénéfices pour la navigation (ou l'exploration) [29].

En considérant les visualisations d'informations sur Internet comme des lopins, il est important de considérer qu'il ne sont pas les seuls et qu'ils sont donc sans cesse en compétition avec d'autres. Si l'odeur d'autres lopins

7 Traduction personnelle de l'anglais *information patches.*

8 Traduction personnelle de l'anglais *information scent.*

(comme des sites utilisant des médias plus traditionnels comme le texte) est plus forte pour l'internaute, ou si l'odorat de ce dernier est plus habitué à ces odeurs, le ratio coûts/bénéfices sera sans doute élevé. Plus spécifiquement, si l'internaute a le choix entre un site A de visualisations et un site B qui ne comprend que du texte et s'il n'a pas l'habitude de ‹ lire › des graphiques, son odorat l'amènera sans doute vers le site B puisque le coût d'apprentissage du nouveau média lui paraîtra trop important. Dans ce manuscrit, je nomme ce coût le **coût de la littératie.**

De la même manière, l'aspect souvent générique des visualisations fait qu'elles ne communiquent pas le sujet des données. Il faut donc que l'internaute passe un certain temps à lire différents titres, labels et autres éléments textuels pour comprendre le contexte de la visualisation avant même de pouvoir commencer à extraire de l'information. Ce coût peut réduire l'odeur de l'information et peut réorienter l'internaute vers un autre lopin. Dans ce manuscrit, je nomme ce coût le **coût de l'interprétation du contexte.**

De plus, la visualisation est un média à fort potentiel interactif, ce qui n'est pas le cas de nombreux autres médias sur Internet. Si l'internaute arrive sur un site de visualisations en s'attendant à une interaction *passive* [[164], [245]] et si la visualisation ne fournit pas d'indice sur son interactivité, alors le coût de la découverte ‹ à l'aveugle › des fonctionnalités interactives sera susceptible de mener l'internaute vers un autre lopin. Dans ce manuscrit, je nomme ce coût le **coût de la perception d'interactivité.**

Enfin, si l'internaute parvient à surmonter tout ces coûts, mais qu'il lui manque des connaissances sur le sujet des données, il aura sans doute du mal à formuler des questions initiales pertinentes et intéressantes. Ceci peut alors entraîner un manque de motivation pour explorer les données qui peut entraîner l'internaute vers un autre lopin. Dans ce manuscrit, je nomme ce coût le **coût de la motivation initiale** [à explorer des donner].

En somme, la visualisation d'informations n'est qu'un média parmi d'autres sur Internet. Les audiences occasionnelles n'ont pas forcément l'habitude de ‹ lire › ce média et peuvent en préférer d'autres si ces derniers leur demandent moins d'efforts (c'est-à-dire s'ils sont moins coûteux). Dans ce manuscrit, je considère les coûts identifiés ci-dessus, qui sont une décomposition des coûts de la perception et de l'exploration de van Wijk, comme les quatre obstacles à l'engagement. Mon hypothèse est que d'aider les audiences occasionnelles à les surmonter favorisera l'engagement de ces personnes dans l'exploration personnelle et efficace des données ouvertes.

Pour finir, afin d'opérationnaliser ma question de recherche principale en prenant en considération chacun des ces coût, je propose les quatre questions ci-dessous :

* **Q1:** Comment un designer peut-il savoir si une audience est capable de comprendre différents types de représentations visuelles de données ?
* **Q2 :** Comment designer une visualisation afin qu'elle illustre immédiatement son contexte ou la sémantique des données qu'elle encode ?
* **Q3 :** Les internautes ont-ils l'habitude d'interagir avec des visualisations – spécifiquement quand celles-ci sont intégrées dans des pages qui contiennent aussi du texte – et, si ce n'est pas le cas, comment les designers peuvent-ils aider ces utilisateurs à détecter le potentiel interactif des visualisations d'informations ?
* **Q4 :** L'intégration d'incitations à explorer les données dans le design d'une visualisation peut-elle motiver les audiences occasionnelles à aller au-delà de la simple représentation par défaut ?

Pour répondre à ces questions, et plus généralement au problème de l'engagement des audiences occasionnelles dans l'exploration de données, j'adopte deux approches : une approche *expérimentale* et une approche *design.* Cette complémentarité est l'une des principales originalités de ce travail.

Pour **Q1**, j'adopte une approche expérimentale afin de montrer l'existence du problème de la *littératie en visualisation* et pour mesurer le niveau de compétences des gens. Pour l'instant, je définis grossièrement la littératie en visualisation comme *la capacité à utiliser efficacement et en toute confiance* *des visualisations pour extraire de l'information de données.* Pour **Q2,** j'adopte une approche design pour trouver une manière de communiquer le contexte d'une visualisation. Pour **Q3,** j'adopte d'abord une approche expérimentale pour montrer qu'une majorité d'utilisateurs non avertis ne cherchent pas à interagir avec des visualisations, puis j'adopte une approche design pour trouver des moyens de suggérer leur interactivité. Enfin pour **Q4,** j'adopte une approche expérimentale pour déterminer si certaines conventions de design de visualisations peuvent servir à engager des audiences occasionnelles dans l'exploration de données.

II Mesurer le Coût de la Littératie: Une Méthode Structurée pour Évaluer la Littératie en Visualisation d'Informations

Dans un article posté en avril 2012, Jason Oberholtzer décrit deux graphiques présentant des données historiques, politiques et économiques sur le Portugal [14]. Bien que n'étant pas expert du sujet, il estime que les graphiques lui ont permis de se forger une opinion bien renseignée sur le pays. Il attribue ce sentiment à la simplicité et à l'efficacité des graphiques et conclut en disant : « Voilà la beauté des graphiques. Nous les comprenons tous, n'est-ce pas ? »[9] Mais les comprenons-nous réellement *tous* ? Bien que le nombre de personnes habituées à voir des graphiques continue à croître, il est encore difficile d'estimer si une personne sait interpréter ou lire un graphique ou une visualisation ; et ce **coût de la littératie** peut être une barrière importante à l'engagement des audiences occasionnelles.

II.I Question de Recherche

Cette section résume le Chapter 3 qui reprend un article publié intitulé A principled way of assessing visualization literacy [122]. Elle étudie le **coût de la littératie** et répond à la question suivante :

Q1 : Comment un designer peut-il savoir si une audience est capable de comprendre différents types de représentations visuelles de données ?

Bien que cette question concerne les designers, elle est tout aussi importante pour les chercheurs. D'ailleurs, la motivation initiale de ce travail a été

9 Traduction personnelle de l'anglais.

de trouver un moyen de mesurer la littératie en visualisation afin qu'elle ne soit pas une variable parasite dans d'autres expérimentations.

II.II Contributions Scientifiques

Les contributions de ce travail sont les suivantes*:*

* une définition concrète de la littératie en visualisation,
* une méthode pour:1)évaluer la pertinence d'items de tests de littératie, 2)mesurer les compétences d'une personne, 3) créer des tests rapides et réutilisables pour des représentations graphiques bien connues, et
* une implémentation de quatre tests disponibles en ligne, basés sur cette méthode.

Tout d'abord, je définis la littératie en visualisation de la manière suivante*: la littératie en visualisation est la capacité qu'a un individu de traduire des questions posées au niveau visuel en questions relatives aux données, et de traduire des questions relatives aux données en « requêtes » visuelles.* Il est à noter que cette définition sous-entend la littératie comme une *compétence passerelle.* Pour faire une analogie avec l'alphabétisation, cette compétence passerelle est celle qui permet d'interpréter des chaînes de caractères typographiques comme un ensemble de mots, de phrases et de paragraphes. Ce n'est que lorsque cette compétence est ‹ activée › chez le lecteur qu'il lui est possible de comprendre un texte et d'en faire l'analyse ou la critique.

Ensuite, je décris une méthode structurée pour mesurer la littératie en visualisation que mes collaborateurs et moi-même avons mise au point et je présente une série de tests standardisés que nous avons développés pour des diagrammes linéaires, des diagrammes à barres, et des nuages de points. Notre méthode est basée sur la *Théorie de Réponse aux Items*[10] (IRT) qui est habituellement utilisée dans les sciences de l'éducation, les sciences

10 Traduction personnelle de l'anglais Item Response Theory.

sociales ou la médecine pour mesurer des aptitudes à l'aide de tests et de questionnaires. Ici, nous utilisons l'IRT de deux manières : premièrement, dans une phase de *conception et de calibrage*, nous l'utilisons pour évaluer la pertinence de différents items potentiels pour nos tests ; et deuxièmement, dans une phase d'*évaluation*, nous l'utilisons pour mesurer la capacité qu'ont les personnes à qui nous administrons les tests à extraire de l'information de représentations graphiques. À l'aide de cette méthode, nous développons alors quatre tests (deux pour des digrammes linéaires, un pour des digrammes à barres et un pour des nuages de points) que nous déployons sur *Amazon Mechanical Turk* (AMT). Après calibrage, nos résultats montrent qu'une majorité de *Turkers* ont un niveau de littératie proche de la moyenne attendue par le modèle d'IRT utilisé, voire légèrement en dessous. Nous utilisons alors ces résultats pour raccourcir et simplifier nos tests afin qu'ils soient facilement ré-exploitables sur Internet.

Ces tests bénéficient des propriétés d'IRT qui permettent d'évaluer une capacité non seulement à base de simples scores (ou notes), mais à l'aide d'un modèle qui place la difficulté de chaque élément du test et la capacité de chaque individu testé sur une même échelle continue — dans nôtre cas, l'échelle représente la diversité des capacités à utiliser des représentations graphiques pour extraire de l'information, c'est-à-dire différents niveaux de littératie en visualisation d'informations.

II.III Discussion

Pour ce travail, j'adopte principalement une approche expérimentale. L'objectif est de comprendre comment 1) mesurer la littératie en visualisation d'informations et 2) de concevoir une série de tests pour aider les chercheurs en infovis à détecter si les participants à leurs expérimentations sont capables de comprendre des représentations graphiques de données afin d'éviter des variables parasites. Ainsi, les prérequis pour ces tests sont la rapidité, la fiabilité et la facilité à les administrer. Ceci renforce leur utilité dans d'autres types de situations en dehors de la recherche académique ; par exemple :

* si des designers veulent connaître le niveau de compétences d'une audience ;
* si des enseignants veulent évaluer le niveau de leurs élèves ;
* si des praticiens veulent embaucher des analystes compétents ; ou
* si des hommes politiques cherchent à établir un niveau standard de littératie en visualisation, tout comme le standard d'alphabétisation.

C'est la raison pour laquelle ces tests calibrés sont disponibles à l'adresse http:// peopleviz.gforge.inria.fr/trunk/vLiteracy/home/. Ils peuvent être directement administrés sur le site et ils retournent immédiatement un score d'aptitude, dérivé du modèle d'IRT utilisé pour leur calibrage.

Enfin, bien que déjà utile en l'état, j'estime que ce travail n'est qu'une première étape de recherche sur la littératie en visualisation. C'est pourquoi j'ai aussi rendu le code source des différents tests disponible sur GitHub pour que d'autres puissent s'en emparer et étendre ma méthode à d'autres types de représentations de données.

III Designer le Contexte : Créer le Lien entre la Visualisation d'Informations et la Communication Visuelle

Dans un article posté en octobre 2012, Alberto Cairo reprend un chapitre de son livre The Functional Art dans lequel il identifie certains défis liés au design d'infographies dans les médias [31]. Il déclare qu'une « bonne » infographie doit présenter de l'information de manière claire et structurée, tout en permettant à l'utilisateur de l'explorer comme bon lui semble. Selon lui, une infographie doit être à la fois un outil de communication pour le designer et un outil d'analyse pour le lecteur. Cairo conclut l'article en s'adressant directement aux designers de visualisations de la manière suivante : « il est essentiel de considérer nos priorités en tant que communicants visuels. »[11] Mais quelles sont ces priorités ? Et comment un graphique peut-il communiquer avec un lecteur ? Bien que le langage de la visualisation soit riche, il est aussi très abstrait. Ceci rend difficile la compréhension immédiate du sujet des données ; et ce **coût de l'interprétation du contexte** peut être une autre barrière importante pour les audiences occasionnelles.

III.I Question de Recherche

Cette section résume le Chapter 4 qui reprend un article de poster publié intitulé The CO2 Pollution Map : Lessons Learned from Designing a Visualization that Bridges the Gap between Visual Communication and Information Visualization [123]. Elle étudie le **coût de l'interprétation du contexte** et répond à la question suivante :

11 Traduction personnelle de l'anglais.

Q2 : Comment designer une visualisation afin qu'elle illustre immédiatement son contexte ou la sémantique des données qu'elle encode ?

III.II Contributions Scientifiques

Les contributions de ce travail sont les suivantes :

* une description de l'implémentation de la CO2 Pollution Map ;
* une série de considération relatives à ce design ; et
* un modèle du processus d'appréhension d'une nouvelle visualisation utile au design.

Je présente le design d'une visualisation sur la pollution au CO2 dans différents pays du monde, inspirée par des pratiques de design graphique et de motion design. J'illustre ensuite comment la transposition des considération issues de ces disciplines peut influencer le design de visualisations. Habituellement, l'infovis tente avant tout de répondre à des besoins analytiques, ce qui impose un focus particulier sur la clarté visuelle — afin d'éviter des biais cognitifs dans l'interprétation des données. Toutefois, ce besoin de clarté, qui est souvent synonyme de sobriété (si ce n'est d'austérité graphique), ne renforce pas toujours le message qu'une visualisation tente de véhiculer.

Comme la visualisation est de plus en plus utilisée comme un média (notamment à des fins journalistiques), les designers doivent trouver des moyens d'en optimiser l'aspect communicationnel. Une approche intéressante est d'étudier comment les designers graphiques et d'animations utilisent diverses représentations visuelles pour communiquer de l'information. Essentiellement, leur travail consiste à d'abord capter l'attention du spectateur avant de l'inviter à déchiffrer une image complexe ; ils utilisent des langages visuels construits avec des formes attrayantes qui tentent d'engager le spectateur émotionnellement avec le contenu du message véhiculé, tout en lui donnant des informations contextuelles.

L'approche que je développe ici est inspirée de ces pratiques. Le design du CO2 Pollution Map utilise un système de particules (une technique traditionnellement utilisée en motion design pour simuler de la fumée) pour présenter une métaphore visuelle de la pollution. Cette métaphore sert à 1) encoder les données à l'aide de certaines variables visuelles (dites *encodantes* [183]) et 2) illustrer le sujet à l'aide d'autres variables visuelles (dites *libres* puisqu'elles n'encodent pas de données), de manière à aider les utilisateurs à comprendre rapidement de quoi il s'agit. À partir de ce design, j'élabore enfin un modèle du processus d'appréhension d'une nouvelle visualisation, inspiré d'une part d'un modèle théorique de communication visuelle et d'autre part d'une simplification du processus de visual analytics. Ce processus comprend cinq étapes centrées autour de l'extraction d'informations d'une visualisation. Il combine les deux approches pour illustrer le cheminement type d'un utilisateur confronté à une nouvelle visualisation.

III.III Discussion

Pour ce travail, j'adopte une approche design. L'objectif est de trouver une nouvelle perspective pour le design de visualisations qui prendrait en compte la communication de leur contexte (ou de la sémantique des données encodées) — sans utiliser du texte supplémentaire ou d'autres objets graphiques superflus. Cela nécessite de créer un pont entre les concepts habituellement appliqués en design de visualisations et ceux appliqués en communication visuelle, dont l'application au CO2 Pollution Map semble montrer l'efficacité.

Bien que l'adage souligne l'importance de ne pas juger un livre à sa couverture, celle-ci établit le premier contact entre l'objet et son public : elle attire le lecteur vers son contenu en délivrant juste ce qu'il faut d'informations contextuelles. L'enjeu du design est donc de fournir juste assez de contexte à l'utilisateur pour qu'il ait une idée du contenu de la visualisation (c'est-à-dire du sujet des données) et pour qu'il puisse estimer si ‹ l'objet › vaut la peine d'être exploré. Toutefois, je préconise que cette perspective de communication visuelle doit être considérée avec précaution afin qu'elle n'interfère pas trop avec les potentiels besoins analytiques des utilisateurs.

IV Assister la Perception d'Interactivité : À la Recherche d'Affordances Perçues pour la Visualisation d'Informations

Dans un article posté en août 2012, Lars Grammel décrit onze types de visualisations interactives qui utilisent différents concepts du design d'interaction [1]. Grammel déclare que ces types de visualisations sont un moyen intéressant d'informer une large audience. Il attribue ceci au fait que ces visualisations facilitent une expérience proche du jeu qui est bien plus engageante que celles proposées par un graphique statique ou une vidéo. Mais comment un utilisateur peut-il savoir si une visualisation est interactive ? Bien que l'importance de l'interaction soit reconnue dans la communauté infovis et qu'elle soit souvent mise en avant comme un facteur de succès de certaines visualisations, elle reste néanmoins une composante nouvelle des représentations visuelles de données — et beaucoup de visualisations omettent encore cette composante, même sur le web. Ainsi, dans un contexte comme celui d'un article de journalisme de données, où une visualisation est intégrée avec d'autres médias comme du texte, il semble présomptueux de s'attendre à ce que les internautes sachent qu'ils peuvent ou doivent interagir avec la visualisation pour trouver de l'information. Ce **coût de la perception d'interactivité** peut être un obstacle à l'engagement d'audiences occasionnelles avec le potentiel interactif de la visualisation d'informations et donc avec l'exploration des données.

IV.I Question de Recherche

Cette section résume le Chapter 5 qui reprend un article publié intitulé Suggested Interactivity: Seeking Perceived Affordances for Information Visualization [124]. Elle étudie le **coût de la perception d'interactivité** et répond à la question suivante *:*

Q3 : Les internautes ont-ils l'habitude d'interagir avec des visualisations – spécifiquement quand celles-ci sont intégrées dans des pages qui contiennent aussi du texte – et, si ce n'est pas le cas, comment les designers peuvent-ils aider ces utilisateurs à détecter le potentiel interactif des visualisations d'informations ?

IV.II Contributions Scientifiques

Les contributions de ce travail sont les suivantes :

* une évaluation de l'habitude qu'ont les internautes à interagir avec des visualisations lorsqu'elles sont intégrées dans des articles contenant aussi du texte ;
* un espace de conception pour suggérer l'interactivité ; et
* une évaluation de trois repères d'interactivité différents.

Tout d'abord, je présente une série d'expérimentations contrôlées menées sur AMT qui ont pour objectif de déterminer si les internautes ont l'habitude d'interagir avec des visualisations publiées dans des articles contenant aussi du texte. Les designers de visualisations ont souvent tendance à considérer les visualisations comme des artefacts isolés que des utilisateurs en toute connaissance de cause viennent consulter et explorer volontairement. Toutefois, beaucoup de visualisations sont intégrées dans des pages web qui contiennent d'autres médias comme du texte. Dans ce cas, il n'est pas certain que les internautes soient habitués à interagir avec ces visualisations; et en effet, nos résultats montrent qu'une majorité n'interagit pas.

Ensuite, j'introduis le concept de *Suggestion de l'Interaction* (SI) que je définis de la manière suivante: *la Suggestion de l'Interaction est un ensemble de méthodes utilisées pour indiquer qu'une zone est interactive en attirant sub-*

tilement l'attention de l'utilisateur sans trop affecter sa concentration ni le reste du design de l'interface.

La plupart des fonctionnalités interactives d'un site internet standard utilisent des widgets (des boutons majoritairement) qui reposent sur des métaphores d'objets physiques pour suggérer leur interactivité : ils empruntent des *affordances* à leur analogues physiques. Celles-ci ne sont pas de ‹ vraies › affordances au sens ‹ Gibsonien › [163] puisqu'elles ne supportent pas l'action physique de pointer, cliquer et (potentiellement) déplacer avec une souris, mais elles suggèrent qu'une interaction est possible. Les boutons, par exemple sont représentés avec des effets d'ombres et de biseaux (qui illustrent leur origine mécanique) pour suggérer qu'il est possible d'‹ appuyer › dessus. Bien qu'efficaces dans de nombreux cas, ces analogies sont inapplicables à des objets interactifs plus symboliques ou abstraits — ceux-ci reposent plutôt sur des conventions de design. Les hyperliens, par exemple, utilisent par défaut une variable visuelle (la couleur) et une marque visuelle supplémentaire (un soulignement) pour suggérer qu'ils sont cliquables. Ce design est ‹ lourd › puisqu'il nécessite deux attributs visuels pour signifier une différence unique avec d'autres éléments textuels (l'interactivité) ; et le fait qu'un utilisateur sache qu'il peut cliquer est purement conventionnel.

Les visualisations quant à elles n'ont ni convention, ni analogue physique pour aider à suggérer leur interactivité. De plus, elles sont généralement des *zones graphiques* composées de plusieurs éléments qui peuvent tous être interactifs. Il faut donc trouver un moyen d'intégrer des *repères visuels* faibles dans le design d'une visualisation qui peuvent indiquer l'interactivité de la zone. Pour résoudre ce problème, je présente une étude de 382 sites hautement interactifs fait en HTML5. Mes collaborateurs et moi-même avons mené cette enquête pour déterminer comment des designers d'interaction utilisent différents repères visuels pour suggérer l'interactivité d'éléments graphiques abstraits ou symboliques. Ceci nous a permis de développer un *cadre de conception paramétrique* pour SI, duquel nous avons pu extraire plusieurs considérations pour l'application de repères SI à des visualisations.

Pour finir ce chapitre, je décris l'implémentation de trois repères SI que nous considérons comme étant les plus représentatifs de la diversité de notre cadre de conception et que nous avons ajouté aux graphiques de l'expérimentation initiale. Ceci nous a permis d'évaluer l'efficacité de ces différents repères SI et nos résultats montrent que seul un des trois a réussi à inciter d'avantage de participants à interagir avec les visualisations. J'émets alors l'hypothèse que ceci est dû au fait que ce repère, contrairement aux deux autres, procure du *feedforward*.

IV.III Discussion

Pour ce travail, j'adopte à la fois une approche expérimentale et de design. L'objectif est de 1) savoir si suggérer l'interactivité est nécessaire et 2) trouver des solutions de design pour suggérer l'interactivité d'objets ou de zones graphiques abstraites comme des visualisations dans une page web. Bien qu'un seul des repères évalués se soit montré réellement efficace, il semble nécessaire de suggérer l'interactivité de visualisations. Ici, je me concentre sur des visualisations intégrées dans des articles qui contiennent aussi du texte, mais je pense qu'il en va de même pour des visualisations plus ‹ autonomes ›. Il se peut que d'avantage de personnes s'attendent à ce que ces dernières soient interactives, mais il n'existe pas de réelles conventions pour leur faire comprendre comment ils peuvent interagir avec l'interface.

De plus, bien que nos résultats sur l'efficacité de SI ne soient que préliminaires, nous avons trouvé que les repères les plus subtiles étaient inefficaces. Ainsi, de la même manière que pour les hyperliens, je préconise un design un peu ‹ lourd ›. Toutefois, j'insiste sur le fait que cette directive n'est pas immuable : un travail d'évaluation des différentes possibilités de SI est nécessaire. Celui-ci pourrait par exemple révéler que d'autres repères plus subtiles sont tout aussi efficaces.

Enfin, je considère que le besoin pour SI n'est peut-être lié qu'à une période de transition : les icônes animées ont été initialement nécessaires aux interfaces graphiques, mais elles ont presque toutes disparues du fait de l'évolution des usages. Cependant, je pense qu'il est important d'accompagner les utilisateurs à travers cette phase transitoire afin d'accélérer l'adoption de la visualisation par le grand public.

V Une Tentative pour Fournir des Motivations Initiales à l'Exploration : Utiliser la Narration pour Engager les Audiences Occasionnelles

Dans un podcast de juin 2014, Scott Murray parle de l'engouement pour un nouveau format de visualisation sur Internet où l'auteur de celle-ci raconte d'abord une histoire avec l'appui du graphique avant de laisser les utilisateurs libres d'explorer les données en détail s'ils sont intéressés [105]. Ces visualisations dites *narratives* sont généralement efficaces pour délivrer un message ou pour convaincre une audience ; mais l'usage de techniques narratives dans la visualisation d'informations permet-elle de générer de l'intérêt pour les données ? Et ces techniques peuvent-elles engager les utilisateurs à explorer les données mises à disposition ? Bien que le travail de designers de visualisations et de journalistes soit important pour créer un contexte informationnel, les audiences occasionnelles doivent avoir accès à des outils et doivent être suffisamment motivées pour extraire de l'information à partir données par et pour eux-mêmes. Ce **coût de la motivation initiale** peut être un obstacle important à l'engagement des utilisateurs dans l'exploration des données.

V.I Question de Recherche

Cette section résume le Chapter 6 qui reprend une publication intitulée Storytelling in Information Visualizations: Does it Engage Users to Explore Data? [125] Elle étudie le **coût de la motivation initiale** et répond à la question suivante :

Q4 : L'intégration d'incitations à explorer les données dans le design d'une visualisation peut-elle motiver les audiences occasionnelles à aller au-delà de la simple représentation par défaut ?

V.II Contributions Scientifiques

Les contributions de ce travail sont les suivantes :

* une première étude comportementale à grande échelle de la manière dont les internautes interagissent avec des visualisations de données ; et
* trois évaluations de techniques narratives utilisées pour fournir une motivation initiale à l'exploration des données.

Je présente trois évaluations dans lesquelles la narration est utilisée pour générer de l'intérêt chez l'utilisateur ; elle fournit des observations initiales et pose des questions ouvertes sur des thématiques traitées par les données. Ces évaluations ont été menées dans un contexte « réel » de presse en ligne (en collaboration avec le groupe Mediapart) et ont pour objectif de comparer le comportement d'internautes confrontés à des visualisations *exploratoires* — c'est-à-dire qui permettent l'exploration des données présentées — en fonction de la présence ou non d'une narration introductive.

De nombreuses visualisations en ligne utilisent des techniques narratives pour expliquer un jeu de données d'une manière simple et efficace. Selon Mike Bostock et Shan Carter, deux designers de visualisations au New York Times, ces graphiques explicatifs sont préférables pour le journalisme de données puisqu'ils ont l'avantage de montrer immédiatement les informations trouvées dans les données [62]. Cependant, ces graphiques sont généralement peu interactifs et limitent le potentiel d'exploration. Or par définition, la visualisation d'informations est interactive est exploratoire. De fait, trouver des moyens pour rendre les graphiques exploratoires plus accessibles et engageants pour le grand public est important, puisque si les données ouvertes doivent réellement rendre les citoyens plus conscients du monde qui les entoure, alors ces personnes doivent être munis d'outils adéquats pour établir leur propre compréhension des données — pas seulement celle fournie par des journalistes dans des articles écrits d'un certain point de vue et suivant une ligne éditoriale particulière. Ici, mes collaborateurs et moi-même explorons le potentiel de la narration pour générer cet engagement souhaité des utilisateurs. Il est à noter que par engagement, nous entendons spécifiquement l'investissement d'un utilisateur dans l'exploration d'une visualisation. Nos résultats montrent que de proposer des questions initiales aux internautes n'est pas suffisant pour les motiver à passer à l'exploration de données.

V.III Discussion

Pour ce travail, j'adopte un approche expérimentale. L'objectif est de 1) comprendre le comportement des internautes face à des visualisations de données en ligne et 2) voir si l'ajout d'éléments narratifs à une visualisation peut les engager à explorer les données. Bien que nos résultats soient négatifs, les trois visualisations utilisées pour l'étude ont connu un certain succès*:* elles ont reçu de nombreuses visites, ont été reprises par différent sites d'information et ont généré des discussions et des débats intéressants sur Internet. Ceci semble indiquer qu'elles sont engageantes d'une certaine manière, bien que ce ne soit pas la manière attendue.

Ainsi, nos résultats doivent être considérés avec précaution puisqu'il ne décrivent l'engagement qu'à un niveau comportemental et focalisé sur l'exploration de données. De plus, cette approche quantitative est peu commune, ce qui la rend difficile à comparer avec d'autres études plus qualitatives menées sur le même sujet — ces dernières reposent généralement sur des observations ou des analyses de commentaires. J'espère donc que cette étude, ainsi que les données comportementales que nous avons collectées, contribueront à établir une mesure de référence pour l'évaluation d'autres stratégies de design visant à augmenter les motivations initiales à l'exploration.

VI Conclusion

Cette section résume le Chapter 7 qui conclut ce manuscrit. Elle aborde d'abord certaines limitations liées à la définition proposée pour les audiences auxquelles sont destinées les visualisation de données ouvertes sur Internet. Ensuite, elle revient sur ma question de recherche principale et confronte l'hypothèse générale aux résultats obtenus. Enfin, bien qu'elles ne soient pas présentées ici, le manuscrit original en anglais propose des perspectives de recherche future sur les différents sous-coûts de la perception et de l'exploration, ainsi que sur la mesure de l'engagement.

VI.I Définir l'Audience des Visualisation de Données Ouvertes

Un des défis majeurs de ce travail a été de définir l'audience des visualisations de données ouvertes : c'est un facteur important du design centré utilisateur car cela permet de prendre des décisions appropriées pour le design. Dans ce manuscrit, je considère cette audience comme une audience occasionnelle que je définis comme des internautes qui ne sont que sporadiquement confrontés à des visualisations d'informations et qui n'ont de ce fait ni une grande connaissance du domaine ou du sujet des données, ni des systèmes de visualisation. Toutefois, la connaissance qu'a une personne d'un domaine ou d'un sujet particulier est difficile à mesurer, surtout sur Internet, et je souligne donc la généralisation que tend à faire cette définition. Les coûts de la perception et de l'exploration peuvent varier grandement au sein de cette audience et je pense qu'établir des recommandations strictes pour une telle audience est compliqué, voire inapproprié.

Habituellement, les designers se concentrent sur des audiences relativement bien identifiées avec lesquelles ils peuvent interagir et itérer sur des choix de design pour mieux prendre en compte les besoins des utilisateurs. Ces audiences sont généralement catégorisées à l'aide de stratégies de segmentation d'un marché ou d'indicateur psychologiques. Bien qu'une telle approche pourrait se révéler intéressante pour identifier les différentes audiences de la visualisation de données ouvertes, je pense que cela nécessiterait des études de marché très spécifiques (ou des études d'utilisateurs) en fonction de chaque type et de chaque domaine traité par ces données. Par exemple, les audiences intéressées par les données ouvertes peuvent aller d'activistes ou décideurs politiques (qui seront sans doute intéressés par certains jeux de données et pas du tout par d'autres) jusqu'à de simples individus curieux de la société et du monde qui les entoure. Au sein de ces groupes, certaines personnes peuvent aussi être intéressées par le média qu'est la visualisation alors que d'autres pas; et elles peuvent vouloir établir leur propre compréhension et perspective sur les données, alors que d'autres souhaiteront simplement avoir une vue d'ensemble de l'opinion publique.

En somme, je pense qu'il est important de comprendre pourquoi et comment une personne entre en interaction avec une visualisation de données ouvertes sur Internet.

VI.II Engager les Audiences Occasionnelles

Avec ces limitations en tête, je reviens sur la question de recherche principale de ma thèse et sur mon hypothèse principale.

Tout d'abord, concernant la manière dont les coûts de la perception et de l'exploration peuvent limiter l'engagement des internautes, les différentes expérimentations que je présente dans ce manuscrit montrent comment les sous-coûts de Ce peuvent représenter des obstacles pour les utilisateurs non-informés ou inexpérimentés. L'évaluation du niveau de littératie des Turkers présentée dans le Chapter 3 dévoile que certains ont du mal à comprendre les visualisations les plus simples et les plus répandues (comme les diagrammes linéaires et à barres) ce qui laisse penser que ces personnes auront beaucoup de difficultés à interpréter des visualisation plus complexes ou moins conventionnelles (comme des *matrices d'adjacence* ou des *treemaps*). Ceci confirme que le **coût de la littératie** est un prérequis à l'engagement, puisqu'il peut empêcher les utilisateurs de comprendre ce qu'ils regardent. Si ce coût n'est pas réglé, il ne peut y avoir de réel engagement, c'est-à-dire un engagement avec le contenu et pas seulement avec la ‹ jolie image ›. De la même manière, les études préliminaires sur la perception de l'interactivité présentées dans le Chapter 5 montrent qu'une majorité de participants ne cherchent pas interagir avec des graphiques, ce qui veut dire qu'ils ne cherchent pas de prime abord à en extraire de l'information. Toutefois, ceux qui découvrent que la visualisation est interactive optent ensuite pour l'usage de ce média. Ceci confirme que le **coût de la perception de l'interactivité** peut nuire à l'engagement des utilisateurs, même si ceux-ci ont le niveau de littératie requis et qu'ils sont à priori prêts à se servir des visualisations pour trouver de l'information. Si ce coût est réglé, les utilisateurs peuvent s'engager dans l'exploration efficace de visualisations. Toutefois, l'exploration ne peut réellement intervenir que si les utilisateurs ont des questions claires en tête et qu'ils comprennent le contexte de la visualisation (c'est-à-dire le sujet des données). En effet, l'évaluation des techniques narratives présentée dans le Chapter 6 montre que même en introduisant des questions et des observations dans la narration, la majorité des utilisateurs n'ont pas (ou peu) pu régler le **coût de la motivation initiale.** De plus, bien qu'une étude plus fine de l'impact du **coût de l'interprétation du contexte** soit nécessaire, je pense qu'il faut qu'un utilisateur soit préalablement attiré par certains éléments qui introduisent le contexte d'une visualisation pour aller au-delà de la simple ‹ jolie image ›.

Ensuite, afin de remédier aux limitations induites par les coûts de la perception et de l'exploration, l'utilisation des différents cadres de conception proposés dans ce manuscrit montrent des résultats encourageants quant à la réduction des **coûts de l'interprétation du contexte** et **de la perception de l'interactivité.** Par ailleurs, bien que les techniques narratives évaluées dans le Chapter 6 ne se soient pas avérées aussi efficaces que je l'espérais, je ne rejette pas pour autant mon hypothèse principale. Les différentes visualisations utilisées ont générées des discussions et des débats intéressants sur les différents sites où elles ont été répertoriées, ce que je considère comme étant un succès. Une des raisons pour lesquelles les gouvernements ouvrent leurs données est d'enrichir le débat public et malgré le manque d'engagement dans l'exploration des données elles-mêmes, les utilisateurs de ces visualisations ont entrepris une certaine forme d'analyse sociale de ces données à travers la conversation. Ceci soulève une question importante qui n'a pas encore été considérée par la communauté infovis à savoir si la visualisation d'informations pour le grand public doit être considérée comme un outil d'analyse, comme un nouveau média interactif, ou simplement comme un objet intermédiaire social qui favorise la discussion et le débat. Répondre à cette question pourrait aider à mieux définir le concept d'engagement avec la visualisation sur Internet, notamment à savoir s'il s'agit d'un engagement dans l'exploration (comme considéré dans ce manuscrit), à la lecture d'une opinion (par exemple le designer ou un journaliste), ou dans des interactions sociales.

En somme, je pense que la visualisation a un grand potentiel pour engager les citoyens dans l'exploration, l'analyse et la compréhension des données ouvertes. Toutefois, il est important de prêter une attention particulière au type d'engagement souhaité au moment de la conception de ces visualisations, puisqu'il semble qu'elles peuvent en engendrer différents types.

VI.III Dernière Remarque Personnelle

Pour conclure, ce travail m'a convaincu de l'importance de s'assurer en premier lieu que les internautes comprennent l'intérêt de la visualisation et sachent utiliser des représentations visuelles pour interpréter et analyser des données. Je pense d'ailleurs que cet enjeu s'étend bien au-delà du domaine de la visualisation : il s'agit de trouver des moyens d'accélérer le processus d'apprentissage des nouveaux médias afin de s'assurer que les citoyens comprennent les nouveaux canaux d'informations. C'est un travail important qui doit être poursuivi et j'espère que ce manuscrit contribuera à la meilleure compréhension des obstacles auxquels peuvent être confrontés les internautes face à des visualisations d'informations et participera à améliorer leur engagement avec le potentiel réel des données ouvertes.

Chapter 1

Introduction

"The effective functioning of a free government [...] depends largely on the force of an informed public opinion. This calls for the widest possible understanding of the quality of government service rendered by all elective or appointed public officials or employees."—**The OPEN Government Act** [16]

In this dissertation, I explore four initial challenges online users may encounter when confronted with *information visualizations* of *open data, i.e.,* data made readily available by governments and public organizations. These data are often advertised as a means for new democracy and have the potential to enlighten public debate. However, making sense of them is a complex matter. This is why visualization is spreading as a new form of media—especially over the web—since it is known to be effective for making data more intelligible. However, most online visualizations today offer very limited interaction possibilities; they simply illustrate specific subsets of open data previously processed by expert analysts or journalists. This limits the potential for personal and social explorations of the data, which limits the amount of unique insights that can be extracted from them. I argue that it is important to engage citizens with more open (or exploratory) visualizations so they can make sense of these data *for* and *amongst* themselves, as I believe this is the best way to truly "empower *the* people." However, this requires helping them overcome the initial challenges they may encounter.

This dissertation is addressed to several audiences. First, it is addressed to *researchers* in information visualization (infovis), as it presents a series of studies and evaluations, which were conducted either to understand the initial challenges online users may encounter (Chapter 3), or to evaluate the effectiveness of certain design choices for engaging these people in the exploration of open data (Chapter 5 and Chapter 6). Second, it is addressed to *designers,* as it introduces different frameworks either for thinking about visualization design (Chapter 4), or for assisting it. This dissertation also discusses ways in which the assessment of online users' initial difficulties, *i.e.,* the evaluation aspect, can help in the process of designing visualizations (Chapter 3 and Chapter 7).

Note that by designers, I essentially refer to *information visualization designers, information designers,* and *graphic designers.* Coming from a graphic design background, I believe there are many bridges to gap between these similar, yet distant disciplines. Information visualization design can learn from the concepts and principles behind visual communication, while graphic design can learn from the research and knowledge on visual perception developed in the field of Infovis. Likewise, the methodologies used in either disciplines seem transferable from one to the other, strengthening either the design approach to infovis, or the evaluation approach to graphic design. This dissertation is set in this interdisciplinary approach, and I hope it will provide new ground for discussion between both disciplines.

1.1 Context

In November 2009, Steve Mollman posted an article discussing the ever growing amount of data available at the reach of a few clicks, and the arising challenges associated with making sense of these data [206]. To him, "the good news" was that governments and other public organizations were increasingly opening their data online; "the bad news" was that the data were "rather dull." This led to "a booming interest" in data visualization, which in Mollman's words "can transform boring stats into compelling graphical representations explaining our world." While visualization can indeed help explain data—which is why it is becoming an unavoidable medium for communicating about them—information visualization can also help explore data to make sense of them in new ways thanks to different interaction techniques. However, this seems much less acknowledged within the general public.

In this section, I present the general context of this dissertation; I provide an initial understanding of open data, and an overview of the art and science of information visualization—in which I emphasize the scientific and analytic aspects of infovis.

1.1.1 Open Data

Open data are trendy. The number or governmental portals for open data is seven times higher today than in 2010 (56 in 2014 against only 8 in 2010), and the number of datasets available is almost four times higher (431,585 in 2014 against only 138,242 in 2010) [185]. But what exactly are open data? And why such a big rush for releasing them? Here, I introduce the concepts and motivations behind open data, and discuss why, in the current state of things, they do not fully live up to their purpose. I first present *what* open data are, and *why* governments and other publicly funded organizations have taken important measures towards releasing their data. I do not however describe *how* this is done, as collecting, segmenting, structuring, and licensing data is not a concern of this dissertation. I then finish by discussing the current restrictions of open data that may prevent people from engaging with them.

1.1.1.1 *What is Open Data?*

Open data can be summarized as *data that are universally available and accessible for reuse and redistribution, without any legal, technological or social restriction*; they are the "building blocks of open knowledge". Open data are generally cultural, political, scientific, financial, statistical, meteorological, environmental, or transport related.

The online Oxford dictionaries [30] define data as *"things known or assumed as facts, making the basis of reasoning or calculation."* In turn, *reasoning* means *going beyond the information or facts given* [260], by transforming them according to rules (*e.g.*, deductive reasoning), or by making inferences or judgements based on them. As such, data can be considered as *building blocks* of knowledge.

The *Open Definition* [56] defines the following framework for *openness:* "knowledge is open if anyone is free to access, use, modify, and share it—subject, at most, to measures that preserve provenance and openness." This is derived from the *Open Source Definition* [7].

More specifically, a *work* (*i.e.*, an item or a piece of knowledge that is transferred) is open if: 1) it is "available under an Open License" (for more information on licensing, refer to Appendix A); 2) it is "available as a whole at no more than a reasonable one-time reproduction cost, preferably downloadable *via* the Internet without charge;" and 3) it is "provided in a convenient and modifiable form such that there are no unnecessary technological obstacles to the performance of the licensed rights." Thus, open data should be machine-readable, available in bulk, and provided in an open format, *i.e.*, a format that can be processed with at least one free/libre/open-source software tool.

1.1.1.2 *Why Open Data?*

Beyond private efforts, many governments and publicly funded organizations have taken action to open up their data; they promote transparency, the provision of useful resources for social, scientific, and commercial innovations, and public participation. This is attested by the many portals that have been created over last few years (*e.g.*, data. gov, data.gouv.fr, data.gov.uk, data.govt.nz, *etc.*).

“My Administration is committed to creating an unprecedented level of openness in Government. We will work together to ensure the public trust and establish a system of transparency, public participation, and collaboration. Openness will strengthen our democracy and promote efficiency and effectiveness in Government.”—**Barack Obama** [34]

The American *OPEN Government Act 2007* [16], which strengthens the *Freedom of Information Act* (Section 552 of title 5, U.S. Code) [26], promotes “accessibility, accountability, and openness in Government.” It sets a new ground for the public availability of data, by amending part of the Freedom of Information Act (Section 552(e)(3) of title 5, U.S. Code) in the following way: “In addition, each agency shall make the raw statistical data used in its reports available electronically to the public upon request.”

The French *Circulaire du 26 mai 2011* [136] promotes the online access to public information for the sake of transparency of public action, and to encourage digital innovation. It calls for the creation of *Etatlab*, a governmental mission in charge of creating and maintaining a unique inter-ministerial portal for open data named data.gouv.fr.

The OECD’s *Declaration on Access to Research Data from Public Funding, 30 January 2004* [40] promotes international co-operation in science and technology. It recognizes that “fostering broader, open access to and wide use of research data will enhance the quality and productivity of science systems worldwide," and encourages the organization to "take further steps towards proposing Principles and Guidelines on Access to Research Data from Public Funding."

Other governments and organizations have taken similar measures, but listing them all would be too long. Simply note that their motivations are alike, and can be categorized into the three principles mentioned above: *transparency, innovation,* and *public participation.* Ideally, these create a cycle (as illustrated in FIGURE 1.1), in which the data provided for transparency reasons is used by innovators to design new services, which in turn engage citizens to participate in public affairs, generating new data to feed the loop. I briefly describe each of these principles in the following paragraphs.

"Transparency promotes accountability and provides information for citizens about what their Government is doing. Information maintained by the Federal Government is a national asset."—**Barack Obama** [34]

Transparency—As suggested by *Transparency International*'s logo [33], transparency is a strong weapon for fighting against corruption in governments, businesses, and civil societies. Democracy, self-government, and popular sovereignty depend upon "the consent of the governed," who not only have a "need to know," but a fundamental "right to know" [16]. By opening up their data, governments hope to enlighten public debate, to enrich democracy, and to strengthen public trust [[34], [136]].

Innovation—Innovation in the Digital Industry is an important fuel for many Economies, and data is one of its major resources [10]. According to the *Open Data Handbook* [151], the economic value of open data is estimated at "several tens of billions of Euros annually in the EU alone." By opening up their data, governments hope to help developers and businesses invent new uses and services, which will in turn create new jobs, and increase

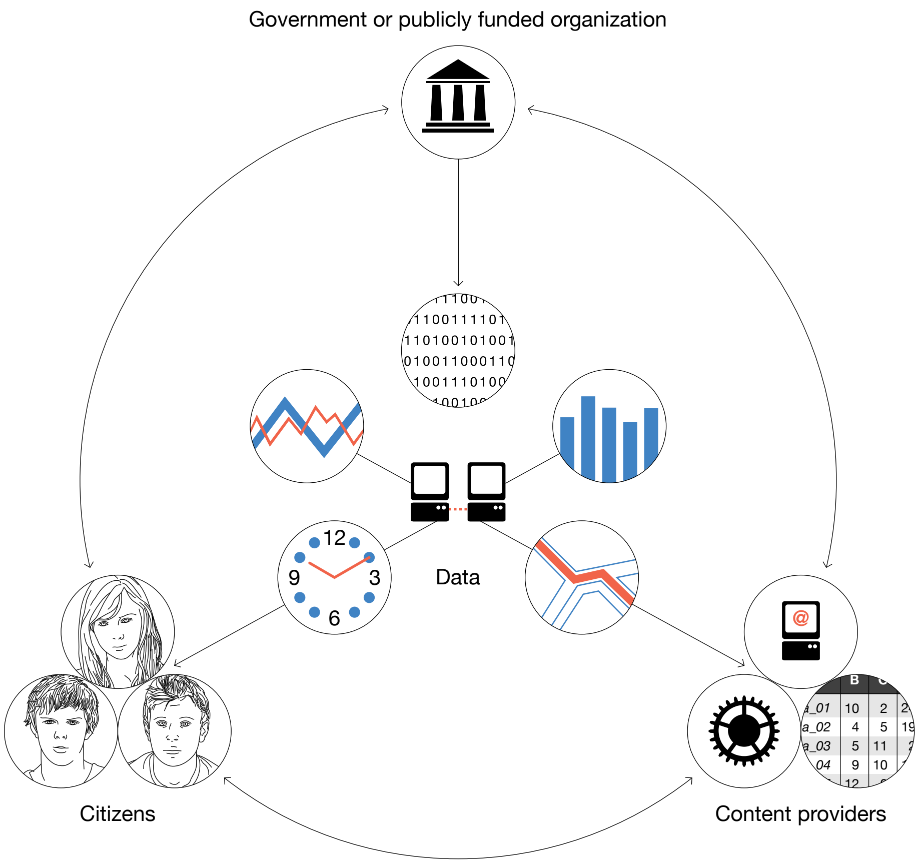


FIGURE 1.1: The open data cycle—personal interpretation of L'Open Data à la Loupe [41].

competitiveness and access to information [136].

Public participation—Simplifying the relation between citizens and public services, and engaging citizens to participate in public affairs are important aspects of the modernization of public action [136]. Collecting feedback from citizens to design better services can also help make governments more efficient and effective [[15] ,[34], [151]]. By opening up their data, governments hope to involve citizens in decision-making, and in the process of governance [40].

1.1.1.3 Discussion

While the motivations listed in Section 1.1.1.2 seem noble, it is difficult to know how well individual citizens can actually make sense of open data. I believe that if these initiatives are to truly enlighten public debate and engage citizens in the process of governance, then open data should be easy to use. Unfortunately, processing and analyzing raw data is a complex matter, which requires specific skills that cannot be expected from everyone. Thus, while the released data respect the terms of the *open framework*, in the sense that they are provided in a convenient and modifiable form that does not suffer any technological restriction *per se* (provided people are equipped with a computer and an Internet access), they do suffer from 'readability' restrictions, since not everyone is able to make sense of them.

Making data 'readable' (or *intelligible*) is usually the job of content providers or of representatives of the news media, *i.e.*, "people or entities that gather information of potential interest to a segment of the public, uses their editorial skills to turn the raw materials into a distinct work, and distributes that work to an audience" [16]. While this dissertation does not intend to criticize the work of these "info-mediaries" [151], I believe they are an extra step between citizens and their governments, which is often driven by a business plan or an editorial line that may limit or orient the information delivered. Once again, to be truly empowered, I argue that *the* people need to make as much sense as possible of open data for and amongst themselves; and to do this, open data must be made accessible in a intelligible form for everyone.

Assuming that information visualization is an effective means to this end, several portals for open data have taken a step toward making their data more intelligible by integrating standard tools that enable people to visualize them (*e.g.*, [[2], [12]]). Similarly, Google has created a service called *Google Public Data Explorer* [24], which aggregates data from different sources, and enables people to explore and relate these data using different visualization tools. In the following subsection, I explain why visualization seems appropriate, although I discuss its possible limitations when it comes to engaging the general public in data-exploration.

1.1.2 Information Visualization

Over twenty five years of research in Infovis has shown that visualization is an effective means for presenting data in an intelligible way. Today it is "deployed against this information onslaught to help us efficiently make sense of and gain insight from the digital deluge" [147]. But what exactly is information visualization? And how can it help people make sense of open data? In this subsection, I provide a broad overview of information visualization, and discuss its benefits and limitations with regard to the general public. I first present *what* it is, and introduce a short history of information graphics. I then describe the value of information visualization, and present the known costs that come with it. After that, I show *how* data can be visualized, and I list common techniques that constitute the 'language' of visualization. I then finish by briefly discussing *why* information visualization seems more appropriate than other data analytics methods for the general public; but I also challenge the idea of an immediate, universal acceptance of this interactive medium.

1.1.2.1 What is Information Visualization?

Information visualization (infovis) is defined as *"the use of computer-supported, interactive, visual representations of abstract data to amplify cognition"* [131], where *abstract* data are data that have no canonical form, *i.e.*, no natural way of being depicted. Like many other visual representations, visualizations are useful for understanding and communicating information, as they help increase the number of items one can consider, and foster mental processing of those items which can lead to new deductions, inferences, and discoveries [261]. This can help in analysis and decision-making tasks. However, information visualizations differ from other more symbolic representations (*e.g.*, text/typography), as they are specifically designed to

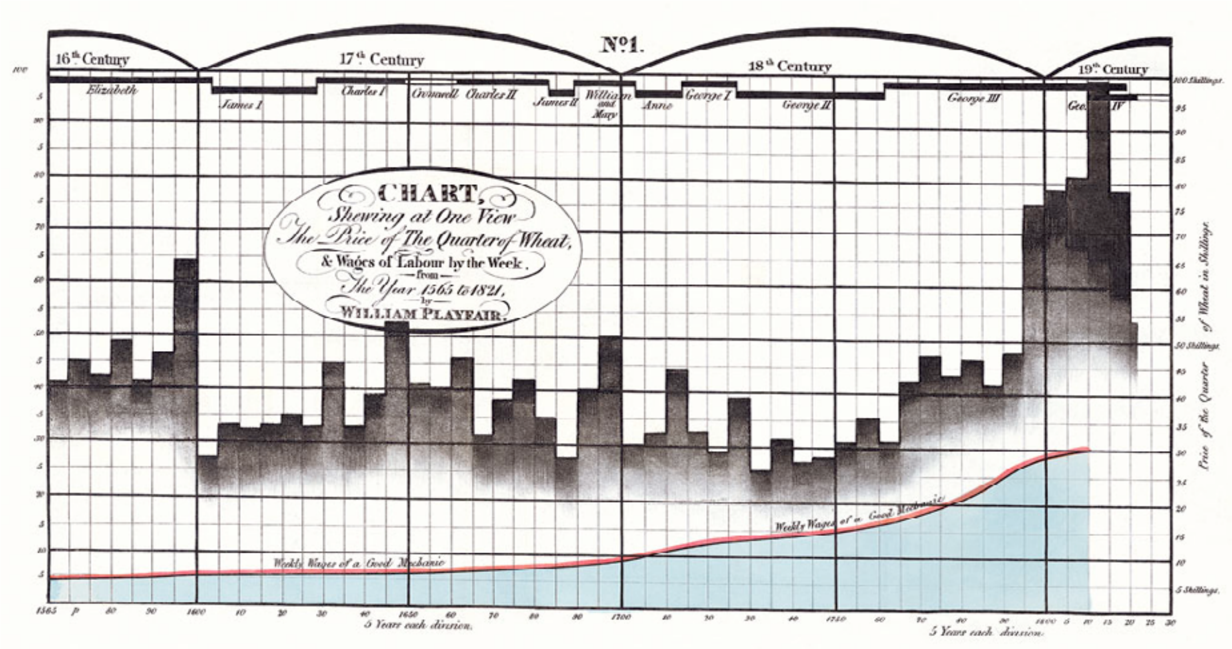


FIGURE 1.2: *Chart Shewing at One View The Price of The Quarter of Wheat, & Wages of Labor by the Week, from the year 1565 to 1821*—William Playfair (1821).

make use of *preattentive* visual properties that can be detected very rapidly (between 200 and 250 milliseconds) and accurately by the low-level human visual system [81]. In addition, information visualizations are interactive, which means that both the data and the visual representation can be manipulated to fit analytic or communication needs.

The origin of infographics—The origin of *information graphics* (infographics) is usually set around the end of the XVIIIth/beginning of the XIXth century, and is marked by the work of Scottish political economist William Playfair; his *Chart Shewing at One View The Price of The Quarter of Wheat, & Wages of Labor by the Week, from the year 1565 to 1821* was one of the most ground-breaking visual representations of its time (FIGURE 1.2). However, its originality was granted with a lot of criticism, to the point that Playfair had to seek justification: “This method has struck several persons as being fallacious, because geometrical measurement has not any relation to money or to time, yet here it is made to represent both. The most familiar and simple answer to this objection is by giving an example. Suppose the money received by a man in trade were all guineas, and that every evening he made a singe pile of all the guineas received during the day, each pile would represent a day, and its height would be proportioned to the receipts of that day; so that by this plain operation, time, proportion, and amount, would all be physically combined. Lineal arithmetic then, it may be averred, is nothing more than those piles of guineas represented on paper, and on a small scale, in which an inch (suppose) represents the thickness of five millions of guineas, as in geography it does the breadth of a river, or any other extent of country” (reported in [[274], pp. 97-98]). Whether this argument managed to convince his contemporaries or not, Playfair led the way for a series of economists, statisticians, and social reformers who would soon use infographics to inform, persuade, and even campaign [21].

The best infographic ever produced—A few decades later, French civil engineer Charles Joseph Minard published one of the most acclaimed infographics of all times: Carte figurative des pertes successives en hommes de l’Armée Française dans la campagne de Russie 1812-1813 (*Map of Napoleon’s disastrous losses suffered during the Russian campaign 1812-1813*—FIGURE 1.3). Étienne-Jules Marey, a contemporary of Minard, wrote about this graphic that “nowhere has the graphical representation of the march of armies reached such brutal eloquence, which [...] seems to challenge the quill of historians” [203](1); and Edward Tufte describes it as “the best graphic ever produced” [255].

The first analytic spot map—Meanwhile, infographics also began to prove their analytic potential for decision-making. In September 1854, English engineer Edmund Cooper created the first *spot map* (FIGURE 1.4),

1 Personal translation from French.

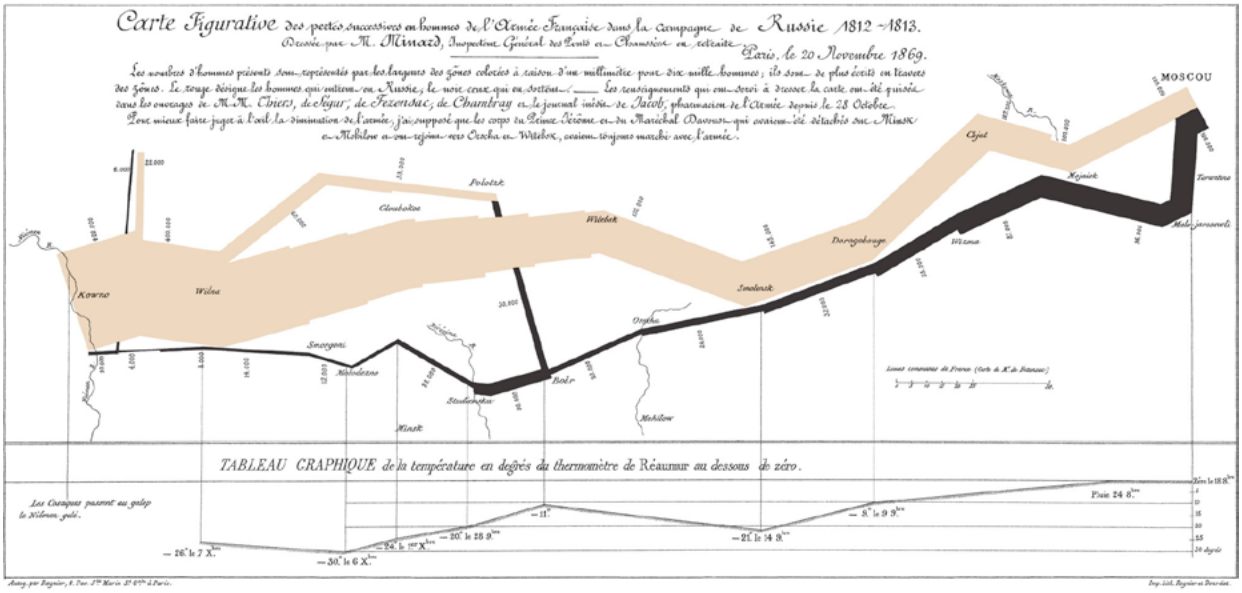


FIGURE 1.3: Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813—Charles Joseph Minard (1869).

while seeking to determine the origin of a cholera outbreak in London. Rumor held that the epidemic was the result of sewer works that had been conducted in the area. By plotting the places where 316 people had succumbed to the disease on a neighborhood map, and comparing this map to the sewer map, Cooper managed to refute the rumor, and determined that it was in fact the "filthy and undrained state of the houses" that was to be blamed [3].

A side-discipline of statistics—By the beginning of the XX$^{\text{th}}$ century, the graphical representation of data had become a side-discipline of statistics. In 1901, French statistician Jacques Bertillon published a set of *Proposals to bring uniformity in the preparation of charts* (in [20]), in order to make their design and interpretation more *easy* and *fruitful*. He suggests several conventions based on six elementary components of diagram design: *points*

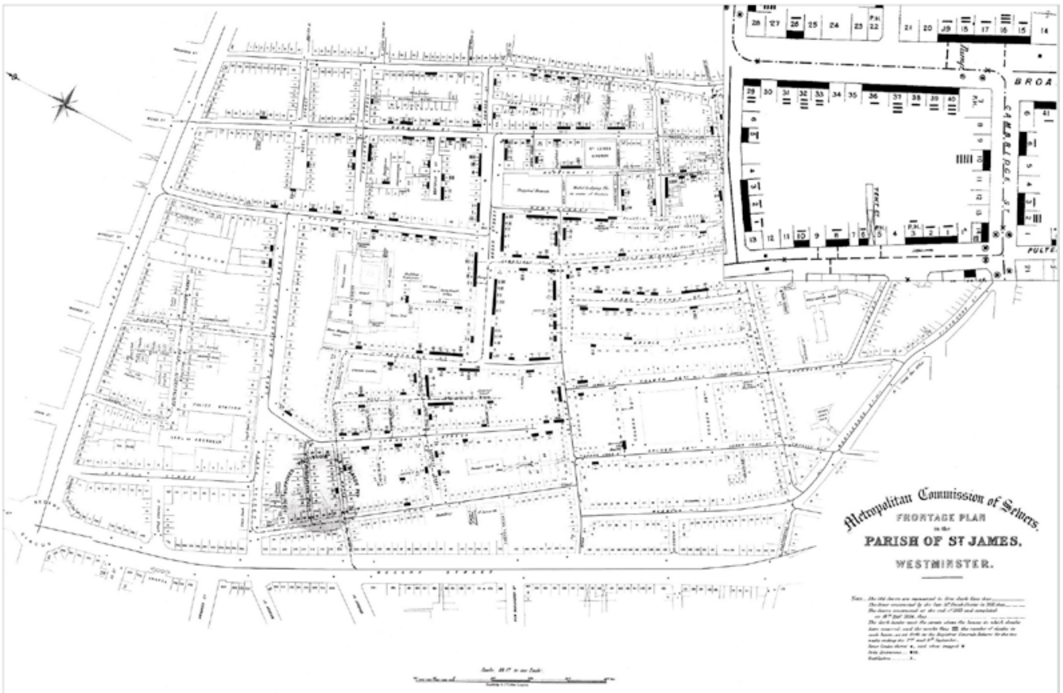


FIGURE 1.4: *Frontage Plan in the Parish of St James, Westminster*—Edmond Cooper, for the Metropolitan Commission of Sewers (1854).

or symbols, lines, surfaces, solid stereograms, colors, and *gradient shades*.

Towards information visualization—In the 1960's, French cartographer and researcher Jacques Bertin proposed a new taxonomy of visual elements, separated into *visual marks* and *visual variables* [118] (FIGURE 1.5). Around the same time, the production of infographics was slowly transferred to computers, as this new technology provided an easier and faster way of processing data and rendering graphics. The use of visualization then spread out to other fields like economics, strategy (military or management), or aviation. However, it mostly remained confined to academic or military applications.

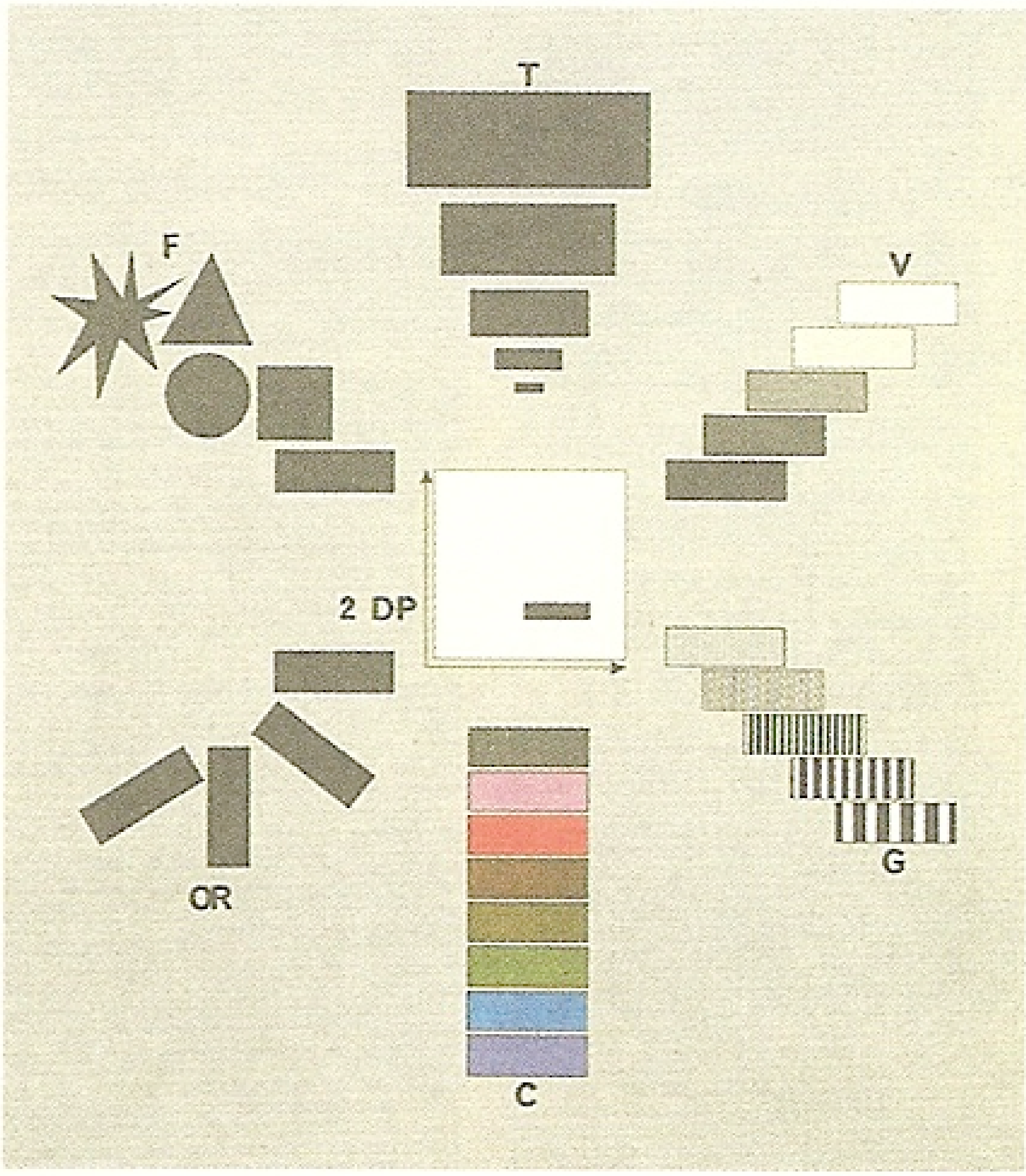


FIGURE 1.5: Bertin's visual variables.

Infovis—It is only with the evolution of mass-media in the 1980's [262] that infographics truly reached out to a large audience. Since then, as a medium for communication and persuasion, infographics have spread in newspapers, on television, and on the Internet, to the point that some are starting to ask: "Where the *hell* did all the infographics come from, anyway?" [11] Meanwhile, with the development of personal computers and business software (*e.g.*, Microsoft Excel), information visualization has become a more widespread tool for data-analysis and decision-making. With this expansion, Infovis finally emerged in the late 1980's [19] as an independent research discipline within the *Computer Graphics* and *Human Computer Interaction* (HCI) communities.

1.1.2.2 *Why Visualize Information?*

The main purpose of visualizing information is to gain insight [212]; it helps amplify the cognitive processes required to analyze data [131] (see Section 1.1.2.1), which in turn helps answer initial questions, and find new ones to ask. Fekete *et al.* [155] have summarized the benefits (or *value*) of infovis in the following list—information visualization can help:

* increase working memory and processing resources available;
* reduce search for information;
* enhance the recognition of patterns;
* enable perceptual inference operations;
* use perceptual attention mechanisms for monitoring; and
* encode information in a manipulable medium.

The value of infovis—The value of infovis is best illustrated by an example. FIGURE 1.6 and FIGURE 1.7 oppose two textual representations and two visual representations of the same dataset. The data is about heart disease related death rates per 100,000 people ('death' column), and

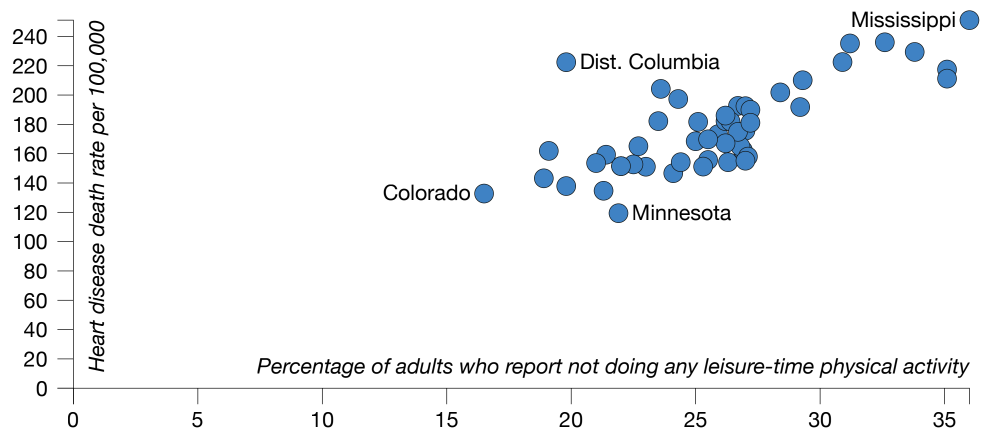
```
state,no_act,reg_act,death
"Alabama",32.6,1.4,236
"Alaska",22,8.9,151.5
"Arizona",24.1,3.2,146.7
"Arkansas",30.9,2.0,222.5
"California",19.1,3.8,161.9
"Colorado",16.5,4.3,132.8
"Connecticut",25.5,3.3,155.7
"Delaware",27,2.6,175.7
"Dist. Columbia",19.8,14.8,222.4
"Florida",26.9,2.2,162.3
"Georgia",26.7,1.8,192.6
"Hawaii",21.3,5.8,134.7
"Idaho",21.4,4.3,159.3
"Illinois",25.1,3.7,181.7
"Indiana",29.2,2.6,191.8
"Iowa",25.9,4.1,173.3
"Kansas",26.8,2.9,164.9
"Kentucky",29.3,2.3,210.1
"Louisianna",33.8,2.4,229.4
"Maine",23,4.3,151.1
"Maryland",26.2,2.6,182.2
"Massachusetts",23.5,5.4,182.2
"Michigan",23.6,2.7,204.2
"Minnesota",21.9,3.5,119.4
"Mississippi",36,1.8,251.1
"Missouri",28.4,2.2,201.8
"Montana",24.4,6.2,154.2
"Nebraska",26.3,3.4,154.2
"Nevada",24.3,2.4,197.3
"New Hampshire",22.5,3.1,152.7
"New Jersey",26.4,3.5,182
"New Mexico",25.3,3.1,151.2
"New York",26.3,6.9,199.9
"North Carolina",26.7,2.0,174.9
"North Dakota",27.1,4.4,158
"Ohio",27,2.6,192.4
"Oklahoma",31.2,2.1,235.2
"Oregon",19.8,6.2,137.9
"Pennsylvania",26.2,4.3,186
"Rhode Island",26.2,4.0,167.1
"South Carolina",27.2,2.3,189.9
"South Dakota",27,4.8,155.2
```

(A) Plain text representation.

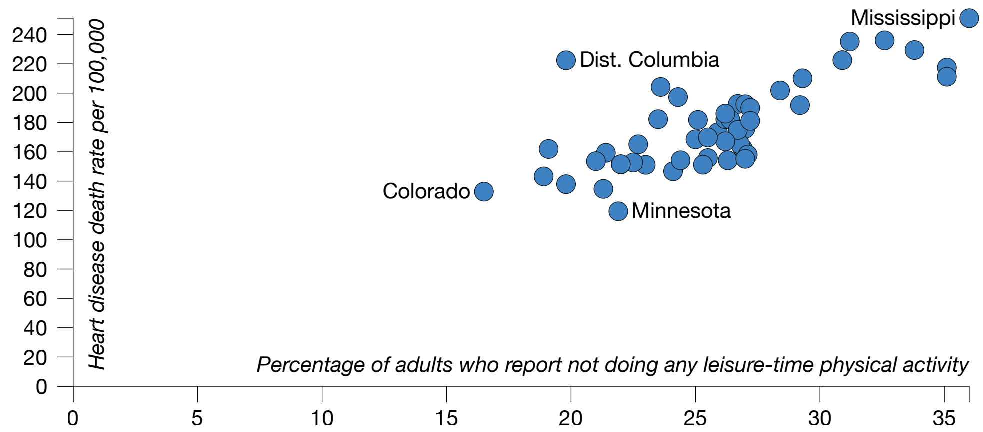
state	no_act	reg_act	death
"Alabama"	32.6	1.4	236
"Alaska"	22	8.9	151.5
"Arizona"	24.1	3.2	146.7
"Arkansas"	30.9	2.0	222.5
"California"	19.1	3.8	161.9
"Colorado"	16.5	4.3	132.8
"Connecticut"	25.5	3.3	155.7
"Delaware"	27	2.6	175.7
"Dist. Columbia"	19.8	14.8	222.4
"Florida"	26.9	2.2	162.3
"Georgia"	26.7	1.8	192.6
"Hawaii"	21.3	5.8	134.7
"Idaho"	21.4	4.3	159.3
"Illinois"	25.1	3.7	181.7
"Indiana"	29.2	2.6	191.8
"Iowa"	25.9	4.1	173.3
"Kansas"	26.8	2.9	164.9
"Kentucky"	29.3	2.3	210.1
"Louisianna"	33.8	2.4	229.4
"Maine"	23	4.3	151.1
"Maryland"	26.2	2.6	182.2
"Massachusetts"	23.5	5.4	182.2
"Michigan"	23.6	2.7	204.2
"Minnesota"	21.9	3.5	119.4
"Mississippi"	36	1.8	251.1
"Missouri"	28.4	2.2	201.8
"Montana"	24.4	6.2	154.2
"Nebraska"	26.3	3.4	154.2
"Nevada"	24.3	2.4	197.3
"New Hampshire"	22.5	3.1	152.7
"New Jersey"	26.4	3.5	182
"New Mexico"	25.3	3.1	151.2
"New York"	26.3	6.9	199.9
"North Carolina"	26.7	2.0	174.9
"North Dakota"	27.1	4.4	158
"Ohio"	27	2.6	192.4
"Oklahoma"	31.2	2.1	235.2
"Oregon"	19.8	6.2	137.9
"Pennsylvania"	26.2	4.3	186
"Rhode Island"	26.2	4.0	167.1
"South Carolina"	27.2	2.3	189.9
"South Dakota"	27	4.8	155.2

(B) Tabular representation.

FIGURE 1.6: Two textual representations of a dataset. Data from [[17], [79]].



(A) A first visual representation showing the relation between heart disease related deaths and percentage of adults who do no physical activity.



(B) A second visual representation showing the relation between heart disease related deaths and percentage of adults who do a regular physical activity.

FIGURE 1.7: Two visual representations of the dataset used in FIGURE 1.6—each dot represents an American state.

the percentages of adults who report either not doing any leisure-time physical activity ('no_act' column), or usually biking or walking to work ('reg_act' column), in the fifty American states and the District of Columbia ('state' column). While it is difficult to extract any kind of information using the 'plain text' view (FIGURE 1.6 (A)), the tabular view (FIGURE 1.6 (B))[2] can help answer simple questions like "Which state has the highest heart disease related death rate?" By scanning the appropriate column, the value can be retrieved quite rapidly. However, this view is limited when it comes to answering more complex questions like "Is there a correlation between heart disease related deaths and absence of physical activity?" and therefore, "Is there an inverse correlation between heart disease related deaths and regular physical activity?" By looking at the two visual representations however (FIGURE 1.7), the answers to these questions become obvious ("yes" and "yes"); and so does the answer to the simpler question. In addition, one can clearly see that the District of Columbia lies outside the two general trends; this may lead to new questions like "Why do adults in the District of Columbia who do regular physical activity still have high heart disease related death rate?" The power of these visual representations lies in the fact that they make use of certain visual properties that can be perceived and processed very rapidly and accurately by the human visual system and brain. However, the limitation here is that two visualizations were needed to represent a single table—and in fact a third visualization could even be produced, comparing the percentages of adults who do no activity with those who do regular activity in every state. This takes up a lot of space, especially when the number of columns (or *dimensions*) in the table grows, and is where interaction becomes useful. Using standard widgets (*e.g.,* drop-down menus), the user could switch the values of the X and Y-axes to see the different relations between each dimension of the dataset in a single view.

2 Note that tables can be considered as visual representations of the data, as they lay the values out in a specific way that helps with certain queries.

“As the eye is the best judge of proportion, being able to estimate it with more quickness and accuracy than any other of our organs, it follows, that wherever relative quantities are in question, a gradual increase or decrease of any revenue, receipt, or expenditure, of money, or other value is to be stated, this mode of representing it is peculiarly applicable; it gives a simple, accurate, and permanent idea, by giving form and shape to a number of separate ideas, which are otherwise abstract and unconnected.”—**William Playfair** [[223], p. x]

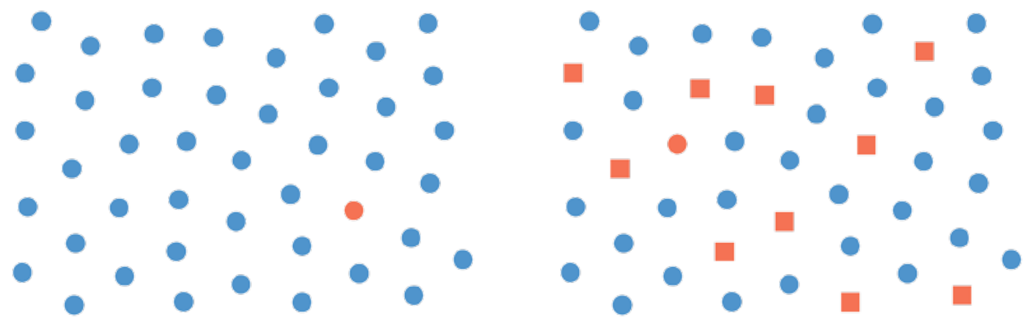
Perception and cognition—An important part of the value of infovis is that it takes into account both the representation of data, and the viewer’s visual and cognitive abilities to process it [170]. Visualizations make use of preattentive *properties* like *hue, curvature, size, intensity, orientation, length,* and *motion* [171] to enhance perceptual inferences, and facilitate information extraction. Several of these visual variables are illustrated in FIGURE 1.5 and FIGURE 1.12. However, while a unique preattentive property can avoid viewers having to focus their attention on a local detail to identify it (FIGURE 1.8 (A)), the combination of these properties can create interference. A *conjunction* occurs when the preattentive properties applied to the local detail are also applied to other distractor elements; this makes the local detail harder to detect (FIGURE 1.8 (B)). In addition, Callaghan has shown that there is a certain hierarchy in the visual processing of preattentive properties [129], meaning that some are favored over others. This is illustrated in FIGURE 1.9, where hue separation is not affected by form randomness, but form separation is affected by hue randomness. Overall, these preattentive properties trigger low-level visual processes that enable viewers to instantly detect targets and boundaries, and count or estimate numbers of visual elements [171]. Meanwhile, other higher-level cognitive processes of perceptual organization, like the *Gestalt laws* (FIG-

URE 1.10), enable viewers to group visual elements by *similarity, proximity, common fate,* and *good continuity* [260]. These favor the detection of patterns and trends in visualizations, like the correlations in FIGURE 1.7.

"Overview first, zoom and filter, then details on demand"
—Ben Shneiderman [243]

Interaction—Another important part of the value of infovis is that it enables users to manipulate both the data and their visual representation (*i.e.,* the *views*), as well as control and share the process and provenance of data explorations [176]. It is quite common that multiple views can be created with a single dataset, and simply laying out all views in the form of a grid (or *matrix*) can be cumbersome, as it requires a lot of space (*e.g.,* in FIGURE 1.7), and can sometimes be visually distracting. In addition, effective data exploration and analysis often requires sorting and filtering operations, as well as moving back-and-forth between contextual views and more detailed views. To perform these data-manipulations, Shneiderman formalized and advocated for the use of *dynamic queries* already twenty years ago [242]. This widget-based approach helps users dynamically formulate and adjust database queries, and is still a standard today. Meanwhile, widgets can also be used to manipulate the way the data is translated into visual form (*e.g.,* in [220]). To perform view-manipulations, different *direct manipulation* techniques [240] have been developed (*e.g., panning and zooming,* or *lasso and rubber band box selections*), which help users select and navigate through the displayed data objects, as well as coordinate and organize views. These interaction possibilities have been categorized in several taxonomies of operations and tasks users can perform with an information visualization (*e.g.,* [[176], [243], [278]]).

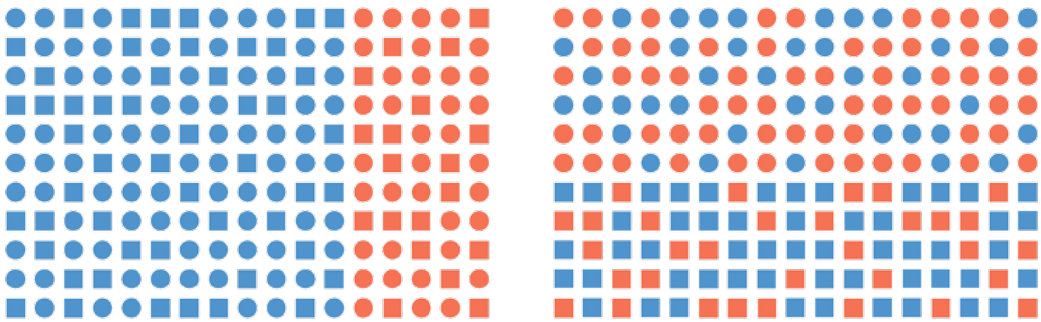
Costs—However, the value of infovis also comes at several costs. van Wijk [263] identifies the four following, which are associated with creating and using an information visualization technique:



(A) A target with a unique preattentive property.

(B) A target with a conjunction of preattentive properties.

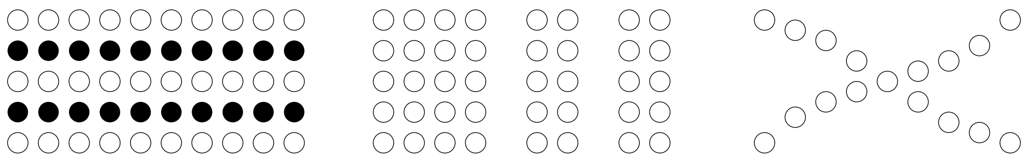
FIGURE 1.8: Typical target detection tasks—personal rendering of [[171], Fig.1].



(A) Boundary distinction by hue.

(B) Boundary distinction by form.

FIGURE 1.9: Typical boundary detection tasks—personal rendering of [[171], Fig.2].



(A) Law of similarity.

(B) Law of proximity.

(C) Law of continuity.

FIGURE 1.10: Three Gestalt laws of grouping—the law of common fate is omitted because dependent on motion.

* Ci—**initial development costs:** the visualization technique needs to be developed and implemented, and new hardware may need to be acquired to do so;
* Cu—**initial costs per user:** the user needs to choose and acquire a visualization technique, learn how to use it, and tailor it to best fit his/her needs;
* Cs—**initial costs per session:** the data need to be acquired, processed, and integrated to the visualization technique; and
* Ce—**perception and exploration costs:** the user needs to understand the visual representation, and learn how to manipulate and explore the underlying data.

Ci and Cs are directly related to creating and setting up a visualization. I also consider Cu to be related to this process, as it assumes a 'software' perspective in which a user must first understand how to use a visualization system (or software) before being able to create a visual representation. Ce however, is related to understanding and effectively using a visual representation to make sense of data.

van Wijk uses these costs in an *economic model* [263] he established to estimate the profitability (in an economic sense) of a visualization technique in order to assess whether it is worthwhile. Essentially, his model boils down to a [return-on-investment - costs] profit formula. If the result is positive, *i.e.*, if the return on investment is higher than the costs, then the visualization technique is worthwhile; otherwise it is not.

1.1.2.3 *How is Information Visualized?*

The reference model for visualizing information is Card *et al.*'s *infovis pipeline* [131] (FIGURE 1.11). The top half presents the 'design' process, starting with raw data and ending with the user; the bottom half presents the user-interaction possibilities at each level. Note that Cs can be associated with the top half, and Cu and Ce with the bottom half—Ci is not associated

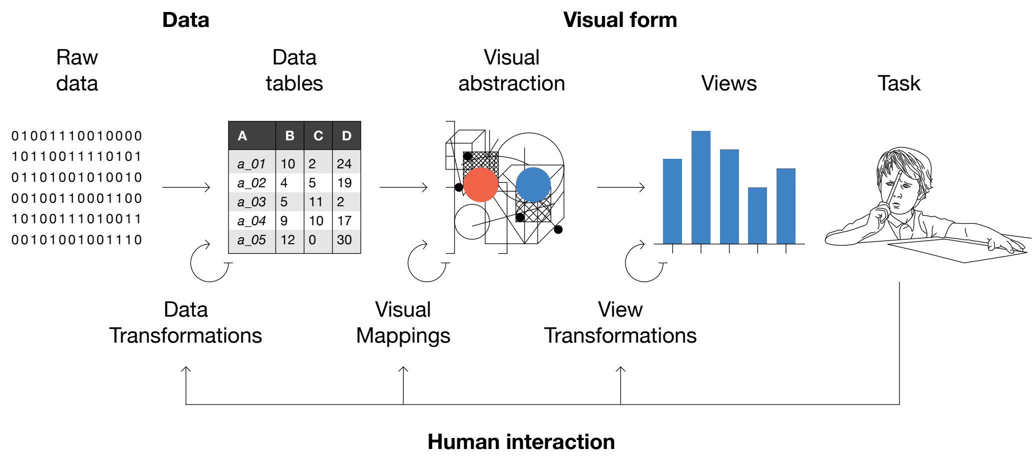
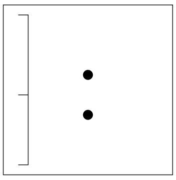


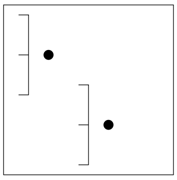
FIGURE 1.11: The infovis pipeline—personal rendering of Card *et al.*'s model [131].

with this model. In addition, while the *Data* segment is an important part of setting up a visualization (and maybe the most time consuming part of the design process), I will not discuss it here, as it is not a direct concern of this dissertation.

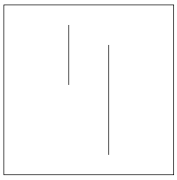
The main act of visualizing takes place in the *Visual form* segment, where data properties are translated (or *mapped*) into visual variables. Cleveland & McGill have evaluated and ranked the accuracy with which viewers can perform quantitative perceptual tasks using these variables [[138], [139]], and Mackinlay has extended and generalized this ranking (see FIGURE 1.13) to other visual variables and perpetual tasks, as well as to other data-types (*i.e.*, ordinal and nominal data) [199]. While these rankings are useful for choosing a visual variable that will best emphasize a specific property of the data, information visualizations are seldom a simple combination of the best performing variables. Besides, not all visual variables in a visualization encode data (*e.g.*, the blue hue of the dots in FIGURE 1.7)—those that do are referred to as *encoding* visual variables, and those that do not as *free* visual variables [183]; and, as discussed above, a poor combination of visual variables can interfere with the desired 'pre-



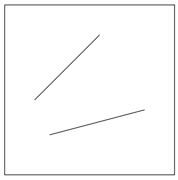
(A) Position (common scale).



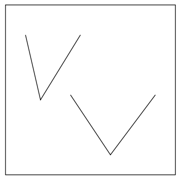
(B) Position (non-aligned scales.)



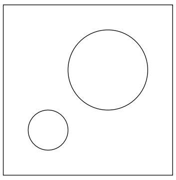
(C) Length.



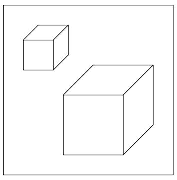
(D) Direction.



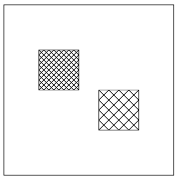
(E) Angle.



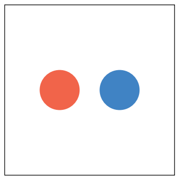
(F) Area.



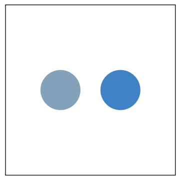
(G) Volume.



(H) Shading or density.



(I) Color hue.



(J) Color saturation.

FIGURE 1.12: Ten visual variables (or elementary perceptual tasks)—adapted and extended from [[138], Fig.1]; curvature and direction are omitted from the original figure, because absent in later work of the authors (*e.g.*, [139]).

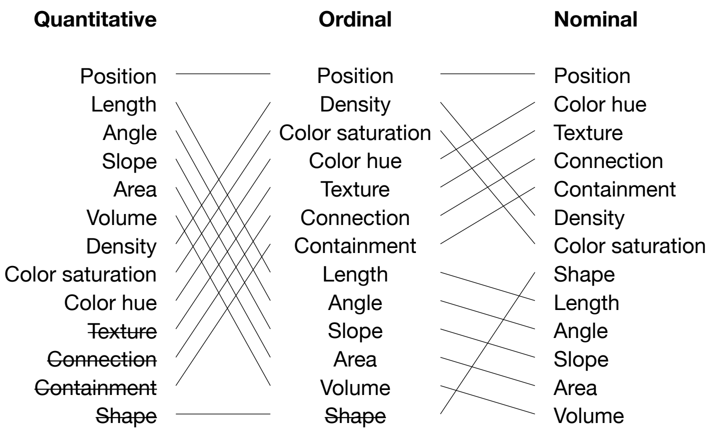


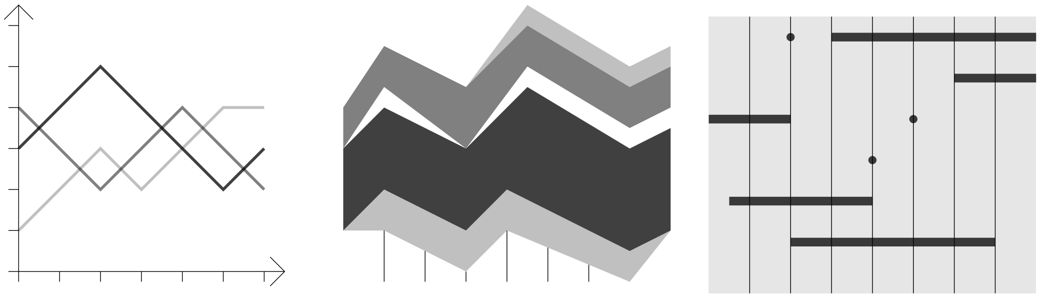
FIGURE 1.13: Ranking of perceptual tasks according to accuracy: barred tasks are not relevant to the data-type—copied from [199].

(A) Azimuthal projection. (B) Mercator projection. (C) Spot map.

FIGURE 1.14: Visualizing space.

attentiveness' of a visualization. Fortunately for design and communication, there are a number of relatively standard visual representations that can be categorized according to the type of data they encode. I briefly present four of these categories in the following paragraphs to inform on the *language* of visualization.

Earth—Be it to delimit territories for living, hunting, or farming, to identify roads or streets, to manage land ownership, or to simply find a restaurant on a smartphone, the visual representation of the Earth has always been at the center of political, economic, social and scientific activities. The Earth is often represented as a *map*, which is a projection of the Globe's quasi-spherical surface (partial or total) on a flat surface (*e.g.*, a piece of paper or a screen). This transformation is not neutral, and several types of projections exist. These can either affect the map visually, by modifying the shape of continents and oceans, and thus, their recognizability (*e.g.*, FIGURE 1.14 (A)); or geometrically, by distorting distances and angles, and thus, their conformity with reality (*e.g.*, FIGURE 1.14 (B)). Therefore, it is important to choose an appropriate projection in accord with the use of the map. On top of these projections, additional visual marks (*e.g.*, circles for spot maps—FIGURE 1.14 (C)—and *bubble chart overlays*) and visual variables (*e.g.*, color hue and saturation for *heatmaps* and *choropleth maps*) can be used to encode other data attributes.

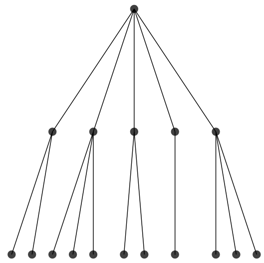


(A) Line graph. (B) Streamgraph. (C) Timeline.

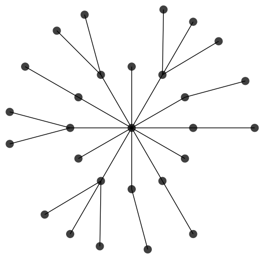
FIGURE 1.15: Visualizing time.

Time—Be it to examine the evolution of global-warming, of the stock market, of biological rhythms, or the history of our institutions, the visual representation of time is useful for understanding trends, and projecting what might come from them. Time is often represented by a series of discrete events projected on a continuous horizontal scale (*i.e.,* a time scale); these *time series* can be encoded as *line graphs* (FIGURE 1.15 (A)), *area charts,* or *streamgraphs* [128] (FIGURE 1.15 (B)). However, not all temporal events can be discretized, and discrete and continuous events may need to be represented together (*e.g.,* in the case of historical events, where individual dates may need to be presented alongside continuous events like a monarch's reign). In this case, a *timeline* (FIGURE 1.15 (C)) is used.

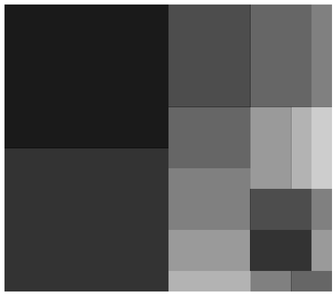
Networks—Be it to identify the links between people in a social network, the connections between regions in the brain, ad-hoc connections between computers and mobile devices, or to describe family descent, the visual representation of networks is useful for understanding the relationships that exist between distinct entities. Networks are often represented as *node-link diagrams,* where the entities (*nodes*) are connected by lines (*links*); these links can be oriented or not. If a network is hierarchical (*e.g.,* a family tree), it has a *tree* structure, and is encoded as a rooted hierarchical node-link diagram (or simply a tree—FIGURE 1.16 (A)), a rooted radial hierarchical node-link diagram (or a *radial tree*—FIGURE 1.16 (B)), or a



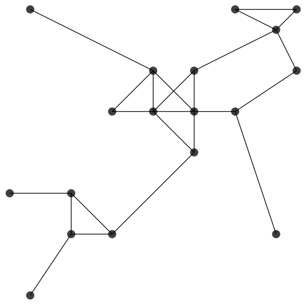
(A) Simple tree.



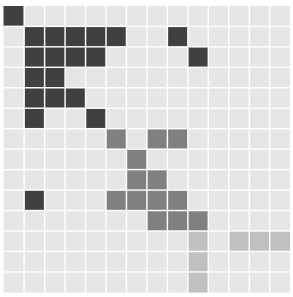
(B) Radial tree.



(C) Treemap.



(D) Simple graph.

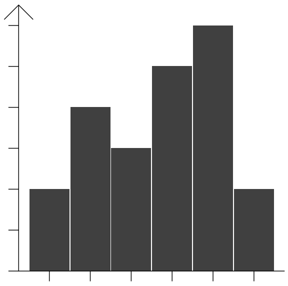


(E) Adjacency matrix.

FIGURE 1.16: Visualizing networks.

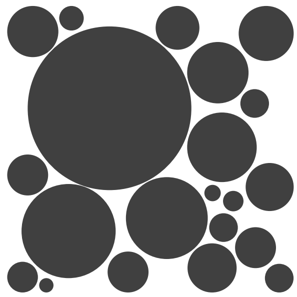
treemap [241] (FIGURE 1.16 (C)). If a network has no hierarchy, it has a *graph* structure, and is encoded as a non-hierarchical node-link diagram (or simply a *graph*—FIGURE 1.16 (D)), or, less frequently, as an *adjacency matrix* (FIGURE 1.16 (E)).

Multivariate data—Be it to budget one's expenses, to compare human development indexes, to seek a correlation between two or more factors, or to compare products across several criteria, the visual representation of multivariate data is useful for analyzing statistical data. Multivariate data can be encoded in different ways, according to the number of dimensions that need to be represented; five categories can be distinguished:

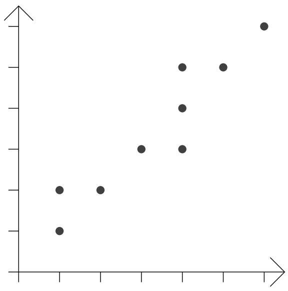


(A) Bar chart.

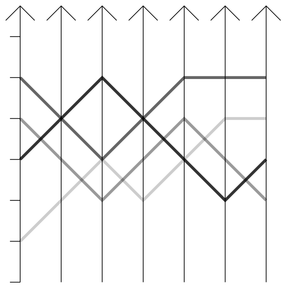
(B) Pie chart.



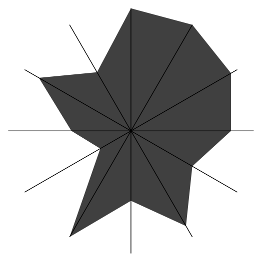
(C) Bubble chart.



(D) Scatterplot.



(E) Parallel coordinates.



(F) Starplot.

FIGURE 1.17: Visualizing multivariate data.

one-dimensional representations (1D), *two-dimensional* representations (2D), *three-dimensional* representations (3D), *n-dimensional* representations (nD), and *high-dimensional* representations. Here, I will only focus on 1D, 2D, and nD representations, as they are the most common. 1D representations can be encoded as *bar charts* (FIGURE 1.17 (A)), *pie charts* (FIGURE 1.17 (B)), *donut charts, bubble charts* (FIGURE 1.17 (C)) or *tag clouds.* 2D representations are usually encoded as *scatterplots* (FIGURE 1.17 (D) and FIGURE 1.7). Finally, nD representations can be encoded as *parallel coordinates* (FIGURE 1.17 (E)), or *starplots* (FIGURE 1.17 (F)).

Note that these categories are neither exhaustive, nor mutually exclusive;

they can be combined in a single visual representation to diversify the information presented. However, they provide an interesting basis for understanding the language of visualization.

1.1.2.4 *Discussion*

While information visualization has a sound theoretical basis that suggests it can be an effective means for making sense of open data, it is not the only way to achieve this goal. Other scientific disciplines like *statistics* and *data-mining* (a subfield of artificial intelligence) also provide techniques and tools for revealing structures in data. However, these techniques are only truly efficient when analytic problems are well-specified and when people have well-defined questions to ask about the data [207]. Most open data analysis problems are not yet specified, and it seems unlikely that people in the general public will know what questions to ask about open data in advance. Moreover, setting up the tools for statistical analysis and data-mining generally requires a high level of programming knowledge, which is not common ground. Finally, visualization also has certain descriptive advantages over these approaches: *Anscombe's Quartet* (see [18]) nicely illustrates how visualization can help see patterns in data that statistical methods cannot describe. Thus, visualization seems overall more appropriate for the general public.

However, I stress that information visualization is not always as engaging and effective as it could be for this audience—otherwise, this dissertation would not be. First, it is unclear whether people are able to understand visualizations as representations of data, or that if they are, they will perceive this medium as being effective and efficient for extracting information. Second, the language of visualization is very abstract and often unappealing. Standard *business graphics,* like bar charts, line graphs, and pie charts—more generally, the kind produced with a few clicks in Microsoft Excel, which are unfortunately most common—are quite dull, and may not intrigue viewers at all. Third, beyond the simple presentation of data, information visualizations are also meant to help people explore data, especially when these are numerous and complex. This requires an awareness of the interactive potential of visualizations, which may not be common ground. Finally fourth, as people may lack expertise or background knowledge on given open datasets, it is possible they will have difficulties formulating interesting questions to trigger an exploratory behavior. This may prevent them from engaging in the process of making sense of data for and amongst themselves, *i.e.*, from going beyond the surface of visualizations, and suggests they may need initial external incentives.

Overall, it seems many people don't know what to do with data, or with visualizations [37]. Several 'visualization as a service' websites have been forced to shut down in recent years due to lack of user-interest (*e.g.*, [[13], [39]]). While this is certainly due—at least in part—to van Wijk's costs associated with creating and using visualizations [263] (see Section 1.1.2.2), I believe many online visualizations suffer from a failure to consider who their targeted audience is. To truly help *democratize* information visualization—and therefore access to open data—I stress the importance of considering *who the* people are, as well as what the context is in which they encounter visualizations.

1.2 Visualization and *the* People

True democratization of information visualization would require enabling *the* people with the possibility to create visualizations, to make sense of the data they chose to explore, to publish their findings within a visualization, and to discuss these findings with others [[180], [267]]. However, this suggests that people already have a general understanding of the purpose and value of information visualization, as well as of the way it 'works;' and that they are willing to spend time and effort learning specific software or programming languages to achieve this. Unfortunately, not everyone is a designer nor a developer. Likewise, not everyone is an expert data-analyst nor a statistician. I personally believe that it is more important, as a first step, to make sure people understand the purpose and know how to use visual representations to make sense of data before attempting to provide them with complex tools and systems for setting up their own visualizations. Therefore, in this dissertation I focus on decomposing, understanding, and designing for overcoming van Wijk's perception and exploration costs (Ce) [263]. The general research question I address is as follows:

How might the perception and exploration costs associated with using an information visualization limit people's engagement in efficient explorations of data, and how might these limitations be remedied?

To better understand these costs, and to operationalize this general research question, it is necessary to find lower-level research questions to address. In this section, I highlight the importance of considering Ce when it comes to the new audiences and contexts in which information visualization is being deployed. I first discuss the shift from traditional infovis audiences towards more casual audiences, and highlight the 'external difficulties' visualization as a medium may encounter on the web. After that, I introduce four *sub-costs* of Ce, using an analogy with the concept of *information interaction,* and with the theory of *information foraging.* I then finish by providing a *takeaway* model that illustrates these sub-costs, and propose four lower-level research questions, which I address in the different chapters of this dissertation.

1.2.1 Casual audiences

"We are seeing a paradigm shift in infovis, away from the specialized user [...] to a more general audience"
—Robert Kosara [36]

The traditional audience of information visualization is composed of workers in fields of high expertise, who have relatively clearly identifiable goals and needs. I qualify these users as *expert* audiences, *i.e.,* people who are confronted with visualizations in a work environment, who have a high degree of domain-knowledge, and who can afford to spend time learning new visualization systems or techniques, provided that these are guaranteed to increase or facilitate productivity. However, citizens who are interested in open data do not necessarily have explicit goals and needs, and cannot be expected to be expert economists, statisticians, social scientists, or data analysts. I qualify these potential users as *casual* audiences, *i.e.,* people who are usually confronted with visualizations only 'on the fly' (if at all), and who do not necessarily have a high degree of domain knowledge, or of infovis systems. These casual audiences cannot be expected to immediately understand the relation between visual encodings and the underlying data, nor the interaction possibilities that an information visualization can provide. Furthermore, it cannot be asserted that they will spend the necessary time learning them, as a leisure activity. This means that Ce is likely to be very high among these audiences.

1.2.2 An online context

"How long will users stay on a Web page before leaving? It's a perennial question, yet the answer has always been the same: Not very long."—**Jakob Nielsen** [28]

Online users generally have a very limited attention span. According to Nielsen, the average time a user spends on a webpage is less than a minute [28]. However, Liu *et al.* [192] have found that web browsing generally shows a *negative aging* effect, meaning that users adopt an initial *screening* behavior, during which they scan a webpage for relevance before deciding to dig for information or not. During this phase, the probability of users leaving the webpage (or *bouncing*) is high, but if it survives this screening process, then the probability decreases, and users are more likely to spend time reading or exploring the page's content. Nevertheless, Nielsen has shown that "realistically," users will only read about 20% of that content [27]. In addition, Liu *et al.* [192] have also found that what they call "less entertaining" content, like Education, Finance, Science, Computers, or Society, is likely to be more harshly screened. Unfortunately, these are the topics of open data. However, I argue that they are interesting, and that casual audiences should be provided with initial help to engage in their comprehension.

To overcome this screening process, commercial websites often include *sticky content, i.e.,* content that retains users' attention, and induces return traffic. Typical sticky contents are weather or news updates, webmail services, chat rooms, or online games. Kominers [188] makes a distinction between *attracting* sticky content, *i.e.,* content that will motivate users to come, or return to a website (*e.g.,* weather or news updates), and *entrapping* sticky content, *i.e.,* content that will motivate users to stay on a website (*e.g.,* webmail services, chat rooms, or online games). However, while these may help generate revenue, they are artificial, as they rarely relate to the content of the website itself.

Finally, a specific problem for online visualizations is that traditional interactions with web-based content are often limited to scrolling up and down a webpage, and clicking through hyperlinks. Interactions with information visualizations are generally more complex and sequential, and occur within a single dynamic page. If a user does not know or understand a visualization, and cannot discover its interactive features, then s/he will likely not perceive the medium as being useful, and will prefer another one—even if s/he is potentially interested in the underlying data. Thus, I emphasize once again the importance of considering Ce for casual audiences, as I expect it will be the primary cause for abandoning a visualization website—such a website should not only be engaging visually (*i.e.*, attractive), but also interactively (*i.e.*, entrapping), so that "people look beyond the surface (*i.e.*, the snapshot visualization), however colorful and pretty it may be" [38].

1.2.3 Four sub-costs of perception and exploration

As mentioned in Section 1.1.2.2, the main purpose of exploring a visualization is to gain insight [212]. Yet, to trigger an exploratory behavior, North suggests that users need to have initial questions; after that, they can go "beyond those initial questions in depth and unexpectedness" [212]. Exploratory behavior is related to question articulation in what Marchionini calls *information seeking* [202], where the process of question articulation, interaction with the query system, and reflective consideration of the outcome is the basis for information tasks. An *information task* is "the manifestation of an information seeker's problem and is what drives information seeking actions." However, exploring an information-rich environment is rarely a single task activity, but rather a process in which each new action is the result of a set of intricate decision points derived from previous actions [250].

Toms describes this information interaction [250] process as a loop that cycles until a satisfactory amount of information is retrieved and integrated. According to her, users can initiate the interaction either by formulating a goal, or simply by deciding to examine a body of information. They select or query a subset of this information, and scan it. When a cue is detected, they stop to consider it, and if it is relevant, they extract and integrate it. Users can then recycle in multiple, nonlinear ways through each step. Thus, information interaction is dependent on both human- and system-factors. The user should have an initial motivation and/or question to answer, and the system should provide clear features for querying and subsetting the provided information. Note that the user must understand how features can be interacted with, and must be able to detect a relevant cue when it is displayed. This should modify the understanding the user has of the information, and should trigger new incentives or questions for the pursuit of the interaction.

Similar to information interaction is the theory of information foraging [222]. Inspired by the strategies animals use when foraging for food, this theory suggests that information seekers are continuously navigating through *information patches,* guided by *information scent.* Patches are regions where information is aggregated, *e.g.,* websites, and scent is a users' estimation of a patch's potential for providing relevant information; scent can be based on personal experience, or on design cues. Nielsen states that the most important concept behind information foraging is a "cost-benefit analysis for navigation" (or exploration) [29].

Considering online information visualizations as information patches, I believe it is important to question the type of information they present, *i.e.,* data, as well as the medium itself, *i.e.,* visualization. In many cases, information seekers on the web stumble across visualizations during a much broader information interaction process that is browsing. This means that *visualization patches* are constantly in competition with other, more traditional media patches, to which information seekers may be more accustomed. If an information seeker is not used to 'reading' from visualizations, s/he is likely to perceive a high cost/benefit ratio in putting the cognitive effort into understanding the graphic. This can reduce information scent, and is likely to lead the information seeker towards another patch. In this dissertation, I refer to this cost as the **literacy cost**.

Nielsen describes Internet users' behavior as *information snacking* [29], where users go online briefly to find "quick answers". Unfortunately, data rarely provide quick answers; they require time and analysis to make sense of, which first demands a minimum understanding of the topic or phenomenon they describe, and of the dimensions they use to describe it. While the abstract nature of the language of visualization is useful for analytic purposes, it creates a generic 'look and feel,' which rarely conveys qualitative information like "what is this about?" at first glance. Thus, before an information seeker can even begin to estimate the cost/benefit ratio of exploring data, s/he must read a series of titles, labels, annotations, *etc.*, to find out what the data are about. This can also reduce information scent, and may make the information seeker leave. In this dissertation, I refer to this cost as the **context-interpretation cost**.

Information visualization can be a highly interactive medium, unlike many others on the web. As mentioned in Section 1.2.2, the interactions a user can perform with a visualization are usually more advanced than the expected scrolling and clicking through hyperlinks. Moreover, visualizations on the web are often embedded with other media like text, and people may simply not be aware that they are interactive. If an information seeker reaches a visualization patch expecting *passive interaction* [[164], [245]], and if the visualization does not provide cues as to its interactivity—especially in a case where it is embedded in text—s/he is likely to perceive a high cost/benefit ratio in attempting to discover *if* the graphic is interactive, and *what* interactions it provides. In this dissertation, I refer to this cost as the **perceived interactivity cost**.

Finally, if an information seeker manages to overcome all of these costs, but does not have sufficient background knowledge about the data or the indicators it uses, s/he may find it hard to articulate initial questions, which may generate a lack of motivation to explore the data. Similarly, if s/he expects to find information upfront, s/he is likely to perceive a high cost/benefit ratio in attempting to dig for it. In this dissertation, I refer to this cost as the **initial incentive** [for exploration] **cost**.

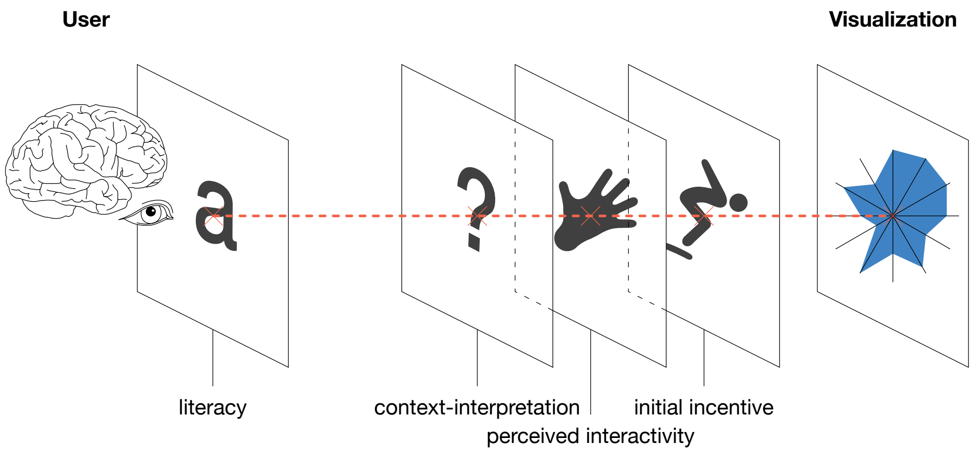


FIGURE 1.18: A simple model of the different sub-costs of Ce.

1.2.4 Takeaway

To summarize, information visualization is but one medium among several others on the web. Casual audiences may not be used to 'reading' visualizations, and may prefer other media if they require less effort to extract the information they are looking for. In addition, Internet users apply a heavy screening process to websites, and are likely to bounce away if they do not find a medium relevant. Thus, is it essential to consider the early perception and exploration costs (Ce) [263] users may encounter when confronted with online information visualizations, in order to retain their attention and engage them in data-explorations. In this dissertation, I propose that Ce can be decomposed into a **literacy cost**, a **context-interpretation cost**, a **perceived interactivity cost**, and an **initial incentive cost**. I hypothesize that these sub-costs are related to necessary, but possibly not exclusive steps casual audiences need to overcome in order to engage with the exploratory potential of information visualizations; and I propose FIGURE 1.18 as a takeaway model, to which I refer throughout this dissertation.

Note that I have set the **literacy cost** on the user's side of the spectrum, as I believe it relates more to personal abilities than to visualization design. In addition, while I have set the other sub-costs in a specific order, I do not claim they are sequential. For example, an information seeker may very well have heard of a visualization (and of the data it presents) before visiting the website it is hosted on. S/he may then enter this model at a step beyond the **context-interpretation cost**. Finally, while these costs may in fact be best represented by linear scales—especially the **literacy cost** for which a user may be familiar with some standard representations (*e.g.,* line graphs and bar charts) but not with more unconventional ones (*e.g.,* adjacency matrices or treemaps)—which would make the cost/benefit ratio or van Wijk's profit formula [263] truly measurable for each case, I simply consider them here as categories of challenges or barriers that a user should overcome in order to engage in efficient explorations of data. My goal is to highlight the existence of these costs and to find appropriate design solutions for reducing them; it is not to extend or refine van Wijk's formula.

To address each of these sub-costs, and to operationalize the general research question presented above, I propose the four following lower-level research questions:

* **Q1**:How can a designer know the level of understanding an audience has of different visual representations of data?
* **Q2**:How can visualizations be designed to help people interpret their context, *i.e.,* the semantic nature of the data they present?
* **Q3**: Do online users have a natural propensity to interact with visualizations—especially when these are embedded with text—and if not, how can we help these people detect the interactive potential of information visualizations?
* **Q4**: Can providing initial incentives for exploration, *i.e.,* external motivations, in the design of visualizations trigger an exploratory behavior in casual audiences, and lead these people to engage in efficient personal data-explorations?

For each of these questions, and to address the overall problem of engaging casual audiences in efficient explorations of data, I adopt two different approaches: an *evaluation* approach and a *design* approach; in one case (for **Q3**), I adopt both. I believe this complementarity makes up one of the main originalities of this dissertation.

For **Q1**, I mainly adopt an evaluation approach for assessing the existence of a *visualization literacy* problem, and for measuring people's ability to interpret visualized data. At this point, I loosely define visualization literacy as *the ability to use common data visualizations (e.g., line graphs or bar charts, which can be found in newspapers, schoolbooks, etc.) to handle information in an effective, efficient, and confident manner.* For **Q2**, I mainly adopt a design approach for finding a way to communicate semantic information about data to a user. For **Q3**, I first adopt an evaluation approach for assessing whether people have a natural tendency to interact with visualizations; and I then adopt a design approach for finding ways to *suggest* their interactivity. Finally, for **Q4**, I mainly adopt an evaluation approach for assessing whether existing design conventions can be used to engage casual audiences in data-exploration.

1.3 Outline

In this final section, I outline the different chapters of this dissertation. After the background chapter (Chapter 2), each focuses on one of the sub-costs of Ce by addressing its associated lower-level research question. The main assumption I have for this work is that helping casual audiences overcome the different sub-costs of Ce will help these people engage in efficient and meaningful explorations of open data.

Chapter 2 presents related work on engagement, and shows how the **literacy cost**, the **context-interpretation cost**, the **perceived interactivity cost**, and the **initial incentive cost** can be articulated around the concept. It then provides an understanding of the constructs behind each sub-cost. After, it describes related work on information visualization *for the people* (or *for the masses*). It then finishes by reviewing several success stories and acknowledged failures of online information visualizations with regard to the sub-costs of Ce.

Chapter 3 focuses on the **literacy cost** by addressing **Q1**. It presents a method for assessing a person's visualization literacy (VL), based on Item Response Theory. It describes the design and evaluation of two visualization literacy tests for line graphs, and presents the extension of the method to bar charts and scatterplots. It then finishes by discussing the reimplementation of these tests for fast, effective, and scalable web-based use, and provides a set of takeaway guidelines for the development of future tests.

Chapter 4 focuses on the **context-interpretation cost** by addressing **Q2**. It describes the design of the *CO2 Pollution Map*, a visualization that takes inspirations from the disciplines of graphic and motion design. This example is used to illustrate how transposing design considerations used in other fields can impact visualization design choices. A framework is then proposed to help rethink visualization design from a visual communication perspective, in order to emphasize the importance of communicating the context of visualizations.

Chapter 5 focuses on the **perceived interactivity cost** by addressing **Q3**. It first shows that a minority of people are naturally inclined to interact with information visualizations when these are embedded in webpages with text. It then introduces the concept of *Suggested Interactivity*, and a *design space* for visual cues that can help users identify abstract interactive features on a webpage. It also presents the design of three SI cues for bar charts, and finishes by evaluating these cues, showing that when a cue provides feedforward, it successfully entices more users to perform interactions.

Chapter 6 focuses on the **initial incentive cost** by addressing **Q4**. It presents the results of three web-based field experiments that assess the impact of using storytelling to 'push' observations, unanswered questions and themes on user-engagement in the exploration of data. In contrast to what was expected, 'pushing' questions does not seem to increase engagement.

Finally, Chapter 7 concludes this dissertation. It provides a general summary of the work presented in the previous chapters, and discusses this with regard to the general research questions set in Section 1.2. It then finishes by detailing perspectives for ongoing and future work on the **literacy cost** and on the **initial incentive cost**, as well as on measuring users' level of engagement with online information visualizations.

Chapter 2

Background

Engagement is an often-cited dimension of the user-experience in information visualization research. Danziger recommends emphasizing the aesthetic and affective appeal of visualizations, to create emotional engagement with their contents; he also mentions that infovis should communicate information to a general audience in "an intuitive and engaging way" [147]. Kosara declares that "we [the infovis community] need to figure out a way to engage people so that they look beyond the surface (*i.e.*, the snapshot visualization), however colorful and pretty it may be?" [38]. Thus, before addressing **Q1**, **Q2**, **Q3**, and **Q4**, it is important to understand how the **literacy cost**, the **context-interpretation cost**, the **perceived interactivity cost**, and the **initial incentive cost** relate in theory to the concept of user-engagement with technological and informational systems. Note that by user-engagement, I specifically mean a user's investment in the exploration of data. Meanwhile, it is also important to understand the various constructs involved in each of these sub-costs, as well as the existing design considerations that may already address them.

This chapter is organized in the following way. It begins by presenting the concept of engagement, highlights how the sub-costs of Ce are theoretically related to it, and discusses how this dimension of the user-experience might be measured. Section 2.2 then presents the different constructs behind of each of these sub-costs, based on related work on graph comprehension, semiotics, perceived affordances, and motivation. Section 2.3 develops on the outreach of information visualization to new audiences, and describes related work on designing information visualizations for *the* people. Finally, Section 2.4 discusses several success stories and acknowledged failures of online visualizations, and attempts to relate these to the **literacy cost**, the **context-interpretation cost**, the **perceived interactivity cost**, and the **initial incentive cost**.

2.1 Engaging Casual Audiences in Data-Exploration using Information Visualizations: A Theoretical Understanding

In this dissertation, I hypothesize that considering the sub-costs of Ce in the design of online information visualizations will help engage casual audiences in efficient and meaningful explorations of open data. I conceptualize engagement from an exploratory point of view, and describe it as an iterative and interactive process, during which a user progressively internalizes his/her motivations for exploring data, thus extending the time spent using the visualization, and the deepness of the exploration.

2.1.1 What is engagement?

Jennings states that "it is the appropriate level of complexity and mystery that will keep the user engaged, but it is immediate positive perceptual judgement of an environment that will entice the user towards exploration and active discovery" [184]. O'Brien & Toms define engagement as "a quality of user experiences with technology that is characterized by challenge, aesthetic and sensory appeal, feedback, novelty, interactivity, perceived control and time, awareness, motivation, interest, and affect" [213]. Attfield *et al.* mention that "user engagement is the emotional, cognitive, and behavioral connection that exists, at any point in time and possibly over time, between a user and a resource" [111]. According to MacCay-Peet *et al.*, "in web applications, user engagement refers to the quality of the user experience emphasizing the positive aspects of the interaction, in particular the phenomena associated with being captivated by the application, and wanting to use it frequently" [218] Finally, Rodden *et al.* describe engagement as the user's level of involvement with a

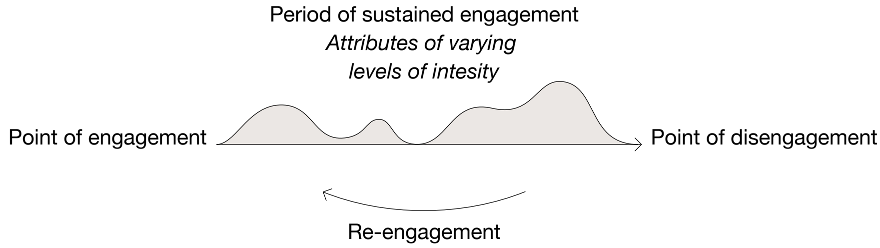


FIGURE 2.1: The four stages of engagement—personal adaptation of O'Brien & Toms' four-stage-model of engagement [213].

product, which translates into the frequency of use, the intensity of use, and the depth of interaction over some period of time [233].

Similarly to Jennings, Tateosian *et al.* [248] consider that engaging visualizations should attract and hold a viewer's attention, by first directing the viewers gaze towards a specific feature of the visualization, and by then encouraging it to linger on that feature. This process of immediate sensory and cognitive appeal leading to entrapment is the basis for O'Brien & Toms' *four-stage-model of engagement* [213], which decomposes the experience into: 1) a point of engagement, 2) a period of sustained engagement, 3) a point of disengagement, and later on 4) a point of re-engagement (FIGURE 2.1). Each of these stages is dependent on several factors, which the authors organize according to McCarthy & Wright's *threads of experience* [204]: the *sensual* thread (Ts), the *emotional* thread (Te), and the *spatio-temporal* thread (Tt). At the point of engagement (and of re-engagement), aesthetics (Ts) trigger positive emotional reactions (Te) and attract users' attention to specific features of the interface; and novel information (Ts) entices interest and motivation to accomplish a task (Te). This immerses users in the "story" of the application (Tt). During the period of sustained engagement, the graphics and information maintain the user's attention and interest, and a rich interface that promote awareness of others and allows for customized views (Ts) creates positive affect like

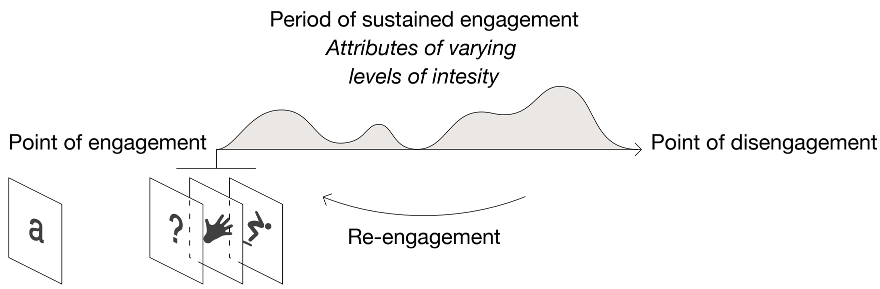


FIGURE 2.2: Articulation of the sub-costs of Ce in relation to the four stages of engagement.

enjoyment, fun, or physiological arousal (Te). This affects users' perception of time, and their feeling of control over the interaction (Tt). Finally, at the point of disengagement, usability issues and/or lack of challenge (Ts) create negative affect like uncertainty, information overload, frustration, or boredom (Te). This can be due to lack of appropriate skills or time, or to interruptions and distractions in the physical environment (Tt). Note that users may also simply disengage because they have succeeded in their tasks, and accomplished their activity (Te). Attfield *et al.* [111] provide a rich summary of the factors related to each of these stages.

In later work, O'Brien & Toms proposed a *structural equation model of engagement* using path analysis to determine the links between aesthetics, novelty, felt involvement, focused attention, perceived usability, and endurability [214]. They found that both aesthetics and novelty can predict felt involvement, and focused attention; aesthetics can also predict perceived usability. Focused attention can then predict felt involvement, which in turn can also predict perceived usability. Finally, both felt involvement and perceived usability can predict endurability.

2.1.2 Articulating the Sub-Costs of Ce in Relation to Engagement

Inspired by O'Brien & Toms' four-stage-model [213], I argue that overcoming the **literacy cost** is a prerequisite for engagement, which can be placed before the point of engagement; all the other sub-costs are related to the point of engagement. A visualization should first attract a user's attention with aesthetically appealing and novel content. This should help focus the user's attention. Of course, this user needs to be able to interpret the visualization as being a representation of data (the **literacy cost**), and the display should provide immediate cues about the topic of the data to immerse him/her in the "story" of the visualization (the **context-interpretation cost**). This should trigger the user's felt-involvement. The visualization should also provide distinct cues about the interactive potential of the display, so that the user can identify whether it is interactive, and how s/he might use it to explore the data (the **perceived interactivity cost**). This should invite him/her to 'try out' the different interactions, and should lead him/her to feel more competent and autonomous, while increasing his/her perceived usability of the visualization. Finally, the visualization should help the user articulate initial questions to trigger the exploration (the **initial incentive cost**), which s/he can then internalize throughout the period of sustained engagement. This articulation of the sub-costs of Ce in relation to engagement is illustrated in FIGURE 2.2, which combines FIGURE 2.1 and FIGURE 1.18.

Note that this proposed model is only an attempt to show how these sub-costs theoretically relate to the concept of engagement. The main purpose of this dissertation is not to validate this model, but rather to explore the different sub-costs themselves.

2.1.3 Measuring Engagement

As the main assumption of this dissertation is that considering each of the sub-costs of Ce in the design of online information visualizations will help casual audiences engage in efficient and meaningful explorations of open data, it is also important to find ways to measure engagement. Regarding the different factors presented in Section 2.1.1, it seems assessing this dimension of the user-experience should rely mainly on qualitative data. For example, O'Brien has used talk-after interviews facilitated through session playback to measure user-engagement in online news interactions [215]. While such data can be relatively easy to acquire in a laboratory environment, it is much less practical, if not impossible to acquire in an online environment. As such, it is necessary to identify appropriate behavioral proxies, which can help describe engagement in analysis or exploration.

Attfield *et al.* have proposed a series of metrics for measuring engagement, including what they call *online behavior* [111]. They suggest that interaction patterns can be instrumental in studying user-engagement. Gotz & Wen have modeled patterns of user behavior in terms of analytic actions [166]. They identify four common patterns: Scan, Flip, Swap, and Drill-Down. A *Scan* pattern describes an iterative set of inspection actions of similar data objects, and indicates a user's intent to compare attributes of these objects. A *Flip* pattern describes an iterative set of changes in filter constraints, and indicates a user's intent to compare multiple sets of the data. A *Swap* pattern describes an iterative set of rearrangements of the order in which dimensions of the data are presented, and indicates a user's intent to find correlations between various dimensions. Finally, a *Drill-Down* pattern describes an iterative set of filter operations on orthogonal dimensions of the data, and indicates a user's intent to narrow the analytic focus to a targeted subset of the data.

From a broader perspective, Rodden *et al.* have proposed a set of user-centered metrics for *Web analytics,* which they categorize in the *HEART* framework, *i.e., Happiness, Engagement, Adoption, Retention,* and *Task success* [233]. Some of these categories rely on attitudinal and subjective measures, which do not fit our present needs. Others however, rely on behavioral metrics and seem adequate for assessing a user's involvement with a webpage. Typically, *Engagement* is measured by the frequency of page views, intensity of each view, and depth of interaction within each session.

In the work presented in Chapter 6 of this dissertation, I have adopted such a quantitative approach to measuring user-engagement with online information visualizations. This is a relatively novel approach, and is another originality of this dissertation, as very few published articles on mass-reaching visualizations have taken behavioral measures into account in their evaluations of success—which I interpret as the level of engagement users have with the visualizations.

2.2 Behind the Sub-Costs of Perception and Exploration: A Theoretical Understanding

Having presented the concept of engagement, discussed how the sub-costs of Ce may relate to it, and identified how it may be measured using behavioral data, it is now important to develop a theoretical understanding of what the sub-costs themselves relate to. In the following sub-sections, I present previous work from several different disciplines, and provide a general understanding of the constructs behind each sub-cost in the following order: 1) the **literacy cost**, 2) the **context-interpretation cost**, 3) the **perceived interactivity cost**, and 4) the **initial incentive cost**.

2.2.1 Behind the Literacy cost: Graph Comprehension

The **literacy cost** is related to people's understanding of how data maps to visual attributes, and vice-versa. If an information seeker is unable to interpret a visualization as a representation of data, or if s/he is unable to understand how the data is depicted by the visualization, s/he is unlikely to find the medium relevant for extracting information. Research into the cognitive processes behind the reading of graphs has been the concern of the area of *graph comprehension* [[196], [221], [228], [239]]. This area studies the specific expectations viewers have from different graph types [259], and has highlighted many differences in the understanding of novices and expert viewers [[160], [196], [252]]. Thus, the challenge of the **literacy cost** is to identify users who may not have the necessary skills for understanding a given visualization, in order to provide them with a more appropriate representation.

Friel *et al.* mention that there generally three kinds of behaviors involved in the comprehension of information presented in a written or symbolic form: translation, interpretation, and extrapolation/interpolation [162]. *Translation* is the act of mentally transforming one form of communication (*e.g.,* textual, visual, *etc.*) into another. *Interpretation* is the act of mentally rearranging and sorting information according to relevance. *Extrapolation/interpolation* is the act of identifying patterns and determining their consequences. Similarly, the OECD's PISA test takes three major aspects of information processing into account: locating, integrating, and generating information [216]. *Locating* tasks require finding a piece of information based on given cues. *Integrating* tasks require aggregating several pieces of information. *Generating* tasks not only require processing given information but also require the examinee to make document-based inferences or to draw on personal knowledge.

Another major aspect of information comprehension is *question posing.* Like for information interaction, but at a much lower-level, Graesser *et al.* [167] claim that question posing is a fundamental component of cognition and that it is a major factor in text comprehension. Indeed, the ability to pose low-level questions, *i.e.,* to identify a series of low-level information extraction tasks to perform, is essential for retrieving information, and for achieving higher-level (or deeper) goals.

Considering question posing for visual representations, Bertin [118] describes three levels on which a graph may be interpreted: elementary, intermediate, and comprehensive. The *elementary* level concerns the simple extraction of information from the data. The *intermediate* level concerns the detection of trends and relationships in the data. The *comprehensive* level concerns the comparison of whole structures and making data and background knowledge based inferences. Similarly, Curcio [146] claims that one can read *from* the data, *between* the data, and *beyond* the data. Friel *et al.* propose a full *Taxonomy of Skills Required for Answering Questions at Each Level,* based on these works [162] (Appendix B).

In addition, several influential models of graph comprehension have been proposed. For example, Pinker [221] describes a three-way interaction between the visual features of a display, processes of perceptual organization, and what he calls the *graph schema,* which directs the search for information in the particular graph. Several other models are similar (see Trickett & Trafton [253]). All involve the following steps:

1. the user has a pre-specified goal to extract a specific piece of information;
2. the user looks at the graph and the graph schema and gestalt processes are activated;
3. the salient features of the graph are encoded,based on these gestalt principles;
4. the user now knows which cognitive/interpretative strategies to use, because the graph is familiar;
5. the user extracts the necessary goal-directed visual chunks;
6. the user may compare two or more visual chunks; and
7. the user extracts the relevant information to satisfy the goal;

Visual *chunking* consists in segmenting a visual display into smaller parts, or chunks [196]. Each *chunk* represents a set of entities that have been grouped according to gestalt principles. Chunks can in turn be subdivided into smaller chunks.

Meanwhile, Shah [239] identifies two cognitive processes that occur during stages 2 through 6 of this model:

- a *top-down* process where the viewer's prior knowledge of semantic content influences data interpretation; and
- a *bottom-up* process where the viewer shifts from perceptual processes to interpretation.

These processes are then interactively applied to different chunks, suggesting that the interpretation process is serial and incremental. However Carpenter & Shah [132] have shown that graph comprehension, and more specifically visual feature encoding, is rather an iterative process than a straight-forward serial process.

Finally, Freedman & Shah [160] relate the top-down and bottom-up processes respectively to a *construction* and an *integration* phase. During the construction phase, the viewer activates prior graphical knowledge, *i.e.*, the graph schema, and domain knowledge to construct a coherent conceptual representation of the available information. During the integration phase, disparate knowledge is activated by reading the graph and is combined to form a coherent representation. These two phases take place in alternating cycles. This suggests that domain knowledge can influence the interpretation of graphs. However, experienced viewers should suffer less influence of both the top-down and bottom-up processes [239].

While the understanding of these higher-level cognitive processes, and the acknowledgement that people have different abilities are important contributions, previous work in graph comprehension has not provided a standardized way of assessing people's ability to interpret visualizations. I posit this is necessary if designers want to create graphics that may suit people's different literacy levels. Therefore, a principled way of testing visualization literacy is required.

2.2.2 Behind the Context-interpretation cost: Signs

The **context-interpretation cost** is related to the ways in which the visual elements that compose the language of visualization communicate what the data are about. If an information seeker is unfamiliar with a dataset and has trouble immediately identifying its topic, s/he is unlikely to seek to understand how it is visually encoded. Research into the use of additional visual features (or *embellishments*) in information visualization has shown that these can have a positive effect on viewers preferences [116], on retention [[116], [120], [121]], and on the effort viewers put into understanding a visualization [179]—which can increase their knowledge and understanding of the data. However, this research is also controversial and has received a lot of critique. Purists like Edward Tufte and Stephen Few qualify such embellishments as *chart junk* [[157], [158], [255]], and highly recommend against their use; they claim that chart junk is an important distraction from data interpretation. Thus, the challenge of the **context-interpretation cost** is to find ways to communicate semantic information about data through visual means (other than text), while staying true to the data and avoiding distraction.

At its most basic level, the language of visualization is the combination of visual marks and variables that form the higher-level representations described in Section 1.1.2.3. However, these representations are by nature very abstract (with the possible exception of maps), and are mainly designed to represent data based on their type and structure, not on their meaning, *i.e.,* on the topic they address.

In the language of *semiotics* [209], visualizations can be considered as signs. A *sign* defines the arbitrary correlation between a *signifier* (*i.e.,* the visual representation) and a *signified* (*i.e.,* the data), along with the *sense* a viewer gets from these attributes. The signifier and the signified need not be directly related, but the sense should be identical for both. For example, seeing a log aflame and reading the word "flame" is not the same thing, but both should convey the same meaning [225].

Signs exist in three forms: icons, indexes, and symbols. An *icon* is a sign in which the signifier and the signified are linked by resemblance, *e.g.,* a map and its territory. An *index* shows a physical connection between signifier and signified, *e.g.,* a crystal glass to signify fragility. Finally, a *symbol* connects the signifier and the signified by means of convention or habit, *e.g.,* the fifty stars that represent fifty states on the American flag. While visualizations are signs, the signified is often the data's structure, and not their topic. For example, an adjacency matrix can signify a graph structure to an experienced viewer (the structure), but it does not signify the nature of the entities that are connected in the graph (the topic). To address this problem, visualization designers often call upon external embellishments, which help convey information about the topic of the data.

Borkin *et al.* have found that adding "human recognizable" objects enhance the memorability of a visualization [121]. They claim that "we are best at remembering 'natural' looking visualizations, as they are similar to scenes, objects, and people." They also state that "making a visualization more memorable means making some part of the visualization 'stick' in the viewers mind. We do not want just any part of the visualization to stick (*e.g.,* the chart junk), but rather we want the most important relevant aspects of the data or trend the author is trying to convey to stick." Bateman *et al.* have also found that such embellishments favor recall, and they have shown that people prefer visualizations with embellishments, and that they find them most attractive [116].

However, violent responses have been formulated against this line of research (*e.g.*, [157]), but I will not enter the chart junk debate here. I will simply evoke Holmes' explanation for the use of images in his charts: "I think [Tufte] missed the point of much that I was trying to do: TIME magazine charts were aimed at lay readers, not unintelligent ones, but busy ones. I knew they'd get the point quicker if they were somehow attracted to the graphic" (reported in [116]). This eloquent quote raises an important question, which is how much designers can or should intervene without orienting the data or the communicated message—thus becoming info-mediaries (see Section 1.1.1.3). Like Holmes, I believe a certain amount of design decisions must be made in accord with the type or topic of the data that is being presented to attract users to the medium. However, such design decision should not affect the amount, quality, or neutrality of the data that is being presented: people should be able to explore and understand unprocessed data for and amongst themselves. Thus, while I argue that the number of steps between governments that provide data and *the* people should be as low as possible, I posit some degree of graphical editing is necessary—at least in an online context where people may not want to spend the necessary time trying to understand what a visualization is about.

Interestingly, this debate on clarity against embellishments and these considerations on the level of intervention (or degree of freedom) a designer should have are not limited to visualization design. Tchischold claims that "a perfect typography is certainly the harshest of the arts [...] for most people a perfect typography does not offer any particular aesthetic appeal, because it is as hard to approach as fine music. The conscience of serving works of quality for a small number of receptive people anonymously, and without expecting any particular recognition is the only reward that a typographer receives for his everlasting learning"[1] [254]. He then argues that typography should be 'transparent' to the uninitiated reader, and that it should not use gratuitous embellishments, so that s/he can focus on the content of the text, rather than on its form. While this may be true for fine type adjustments, it is important to note that typographic choices have an impact on the overall 'look and feel' of a textual document. Choosing to use one font face instead of another gives a particular tone or style to a text. For example, setting a scientific paper in Comic Sans would seem inappropriate. Likewise, setting a butcher's sign in roman capitals would seem odd. Conversely, setting the columns of the Times newspaper in Times font is meaningful.

This analogy with typography reveals that it may be possible to use the 'unexploited' visual attributes of visualizations themselves to convey semantic information about the underlying data. However, it is important these attributes do not interfere with the perception mechanisms used by visualizations. Therefore, I posit the use of free visual variables [183] should be explored to map meta-information, and to bring forth contextual cues in a visualization.

2.2.3 Behind the Perceived interactivity cost: Affordances and Perceived Affordances

The **perceived interactivity cost** is related to people's understanding of *if* and *how* a visualization can be interacted with. If an information seeker is unable to detect that a visualization is interactive, or cannot locate its different interactive features, s/he is unlikely to engage in deep, meaningful explorations. Interface and interaction design practitioners often makes use of specific graphical features for this purpose, generically named *perceived affordances*. The fact that one region of an interface can be interacted with rather than another is not intuitive, but based on a number of commonly shared understandings that such graphical features mean

1 Personal translation from French.

possibility for interaction. Thus, the challenge of the **perceived interactivity cost** is to provide users with identifiable and understandable cues that suggest the possibility for interaction with a visualization.

The term "affordance" was coined by the perceptual psychologist J. J. Gibson to define certain properties of the world that induce action in an organism [163]. More specifically, these properties belong to an artifact or an environment and affordances describe the relationship between them and the organism. Norman introduced the term to the field of design to define the specific attributes of physical artifacts that help people understand the way they ought to be manipulated [211]. However, due to inappropriate use of the term, he was forced to distinguish between 'real' affordances, *i.e.*, the actual properties of an artifact that call for action; and perceived affordances [88], *i.e.*, the perception and understanding a person has of the actions that can be performed with that artifact. He argues that the actions that are perceived as doable in an interface are conventions, or logical and cultural constraints, rather than actual affordances. As he puts it: "it is wrong to argue whether a graphical object on the screen 'affords clicking.' It does. The real question is about the perceived affordance: Does the user perceive that clicking on that location is a meaningful, useful action to perform?" [88] This means that the possibility for interaction with a graphical object should not only be perceived, but also interpreted. Thus, perceived affordances are more a cognitive relationship between a user and an object than a behavioral relationship—described by 'real' affordances. In later work, Norman even introduced the term *signifier* to avoid the confusion [88]. However, to remain consistent with other existing research, I will stick with the term "perceived affordance."

Hartson makes a further distinction between *cognitive, physical, sensory* and *functional* affordances [169] (Appendix C). He emphasizes the importance of dealing with these, and gives the following list of questions to consider when designing an interactive artifact:

* "Is the functionality to which this interaction or artifact gives access useful in achieving user goals through task

performance?" (functional affordance, or purpose of physical affordance);
* "Does the design include clear, understandable cues about how to use the artifact, or about system outcomes if the artifact is a feedback message?" (cognitive affordance);
* "Can users easily sense the visual cues about artifact operation?" (sensory affordance in support of cognitive affordance);
* "Is the artifact easy to manipulate by all users in the target user classes?" (physical affordance); and
* "Can users easily sense the artifact for manipulation?" (sensory affordance in support of physical affordance)

Based on this, Vermeulen *et al.* have proposed the following complete definition of perceived affordances: *"perceived affordances are cognitive affordances that are understandable through well-defined sensory affordances (e.g., a door's handle) and reveal a physical affordance (an action possibility), which is coupled to a functional affordance (the action's purpose). Perceived affordances occur before the user's action and invite them to an appropriate action"* [264].

Finally, Tang *et al.* [247] have even extended Hartson's model with *perceived affective affordances* and *perceived control affordances,* which describe the attributes of an artifact than can trigger or stimulate an emotional reaction in the user, and the attributes of an artifact that give the user a certain level of perceived control over the interaction, respectively.

While affordances are an intricate part of successful interface designs, they often rely on metaphors of real world objects, simulating their physical aspects and properties to suggest a specific action. Unfortunately, interactive visualizations have no real world counterparts that can help suggest their interactivity. Therefore, I posit it is necessary to look at other design conventions used in interface design to suggest the interactivity of specific features, and to see how these may be applied to information visualizations.

2.2.4 Behind the Initial incentive cost: Motivation

The **initial incentive cost** is related to people's immediate motivation for exploring visualized data. If an information seeker does not have initial questions in mind, or if s/he simply does not have an initial motivation for exploring the data, she is unlikely to engage in deep, meaningful information interaction. Understanding and facilitating motivation has been an important research agenda for education. While young students may be intrinsically motivated by learning, most educational activities prescribed in schools are not intrinsically interesting (as mentioned in [235]). Students then rely on some sort of external motivation to achieve their education; and one of the main challenges for teachers is to find ways to help students internalize these motivations. Similarly, while an information seeker may be intrinsically motivated by finding new information, overcoming the fact that it may not be presented upfront may be frustrating; and articulating initial questions to answer in order to start an exploration may not always be easy. Thus, the challenge of the **initial incentive cost** is to provide users with immediate motivations for exploring a visualization, and to find ways to help them internalize these motivations so that they can engage in this exploration process.

Theories on motivation have traditionally made a distinction between intrinsic and extrinsic motivations. *Intrinsic motivation* relates to "the natural human propensity to learn and assimilate" [235]; someone intrinsically motivated will conduct an activity because it is inherently interesting or enjoyable. *Extrinsic motivation* relates to the external regulation of an activity; someone extrinsically motivated will conduct an activity because it leads to a separable outcome. However, Ryan & Deci argue that there are several orientations of motivation [235]. In *Self-Determination Theory* (SDT) [148], they distinguish different types of motivations, based on the reasons or goals that lead people to conduct an activity. In *Organismic Integration Theory,* a sub-theory of SDT, they detail several different forms of extrinsic motivations. Based on this, they propose the taxonomy of human motivation shown in FIGURE 2.3.

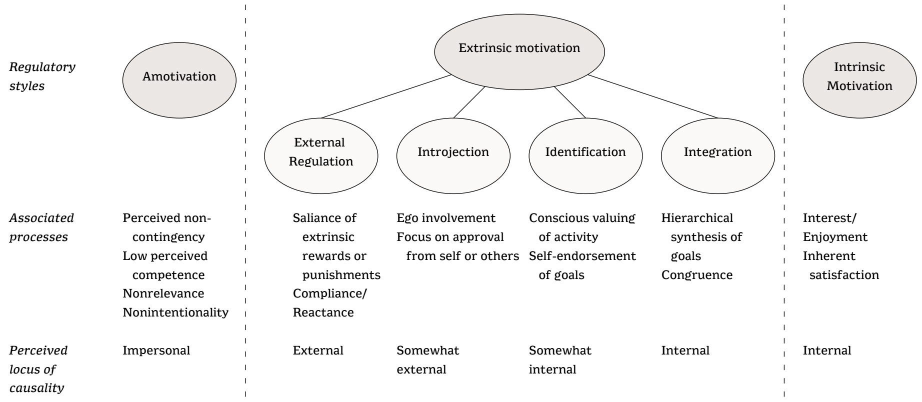


FIGURE 2.3: Human motivation—personal rendering of Ryan & Decy's taxonomy [235].

To the left of the figure is amotivation. *Amotivation* is a state in which someone lacks intention to act; it can result from not valuing an activity, not feeling competent to do it, or not believing it will yield a desired outcome [235]. This is important to consider, as people may lack *self-efficacy* [114] (or perceived competence) regarding interactive visualizations, and may not perceive the value of infovis. In other words, if people are not convinced in their competence, and in the effectiveness of information visualizations, they will likely not engage in exploration at all.

In the middle of the figure is extrinsic motivation. Extrinsic motivation is decomposed into external regulation, introjection, identification, and integration. *External regulation* is the least autonomous form of extrinsic motivation; activities are conducted to satisfy external demands, or to earn external rewards. *Introjection* is a type of internal regulation that is still quite controlling; activities are conducted with the feeling of pressure in order to avoid guilt or anxiety, or to enhance one's ego or pride. *Identification* is more autonomous and self-determined; activities are conducted because they are perceived as personally important. *Integration* is the most autonomous form of extrinsic motivation; activities are conducted congruently with one's other values and needs [235].

Finally, to the right of the figure is intrinsic motivation. As mentioned above, intrinsic motivation is a state in which someone will conduct an activity for its inherent satisfactions—the reward is the activity itself; it is dependent on feelings of self-efficacy and autonomy [235]. Note that each of these stages of internalization do not necessarily create a sequenced continuum; people can initially adopt any of these activity regulations, based on their prior experiences and on situational factors.

Intrinsic motivation is also an important component of *Flow theory*, which is defined as the condition"in which people are so involved in an activity that nothing else seems to matter; the experience itself is so enjoyable that people will do it even at a great cost, for the sheer sake of doing it" [144]. Flow experiences generally have the following characteristics: a merging of action and awareness, a centering of attention, a loss of self-consciousness, a sense of control over the activity, a set of demands for action and clear unambiguous feedback, and autotelism [142]. According to Csikszentmihalyi, Flow experiences "provide opportunities for action which a person can act upon without being bored or worried" [142]. As illustrated by FIGURE 2.4, it occurs when personal skills are well suited to the challenges set by a task or activity. If this activity is too complex, it will lead to anxiety; if it is too simple, it will lead to boredom.

Flow, and more specifically this balance between personal skills and challenges is an important component of game design, which I believe can also be used to engage people in the exploration of data using information visualization. However, this requires having ways to assess people's skills, as well as the level of challenge specific visualization techniques may represent—something that has not yet been addressed. Chen *et al.* [134] have studied how people may enter the state of Flow on the web; they claim that one of the most compelling activities for this is information seeking. However, in light of the examples they provide from their survey, I argue that what they describe as Flow may be more a reflection of participants' internalized extrinsic motivations, than actual intrinsic motivations (a condition for Flow).

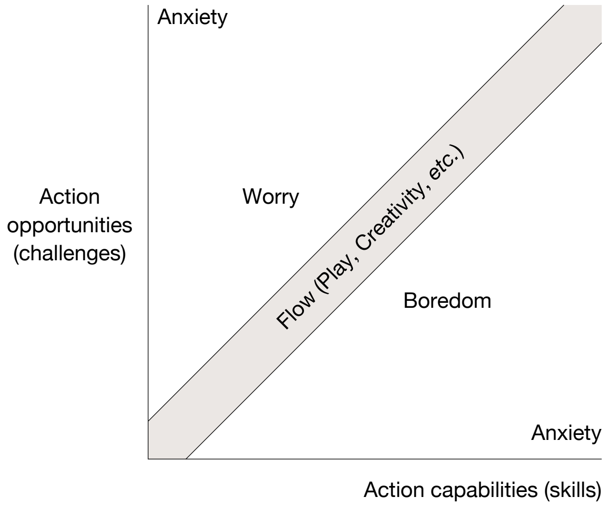


FIGURE 2.4: The Flow state—personal rendering of Csikszentmihalyi's model [144].

To achieve this internalization of motivations, Ryan & Deci suggest providing people with a sense of relatedness, *i.e.*, a sense of belonging to a community [235], and emphasize that they should have the necessary skills to feel competent and autonomous. Overall, this internalization is a desirable goal, and is what should happen during the period of sustained engagement. However, I posit initial questions and cues for exploration should be integrated to the design of visualizations to provide casual audiences with external motivations to trigger the point of engagement.

2.3 Designing Information Visualizations for Casual Audiences: A Practical Understanding

Having shown that each of the sub-costs of Ce has a sound basis in theory, I now turn to the different design considerations that have been suggested for engaging *the* people with information visualization. In this section, I develop on the outreach of visualization to new audiences, and I describe the concept of infovis for *the* people. I present different design dimensions that have been proposed, which I relate to the sub-costs of Ce, and highlight previous work conducted along each of these dimensions. I then finish by discussing several success stories and acknowledged failures of information visualization for *the* people, and in each case, I attempt to assess which sub-costs were addressed and those that were not.

2.3.1 Mass-Reaching Visualizations

As mentioned in Section 1.1.2.1, information graphics have existed in mass-media for the past thirty years now. Utt & Pasternack have traced the origins of journalistic infographics back to the beginnings of *USA Today* in 1982 [262]. They found that by 1991, 71.6% of the newspapers they surveyed had run at least one infographic on an issue's front page; and in 1993, more than half published between three and six infographics every day. However, they also noticed that nearly 90% of infographics in newspapers served only as supplementary information sources to larger bodies of text [262]. This may suggest that at the time they conducted their research, infographics were not considered 'communicative' enough to be published independently.

Mack [198] comments this figure, by declaring that it can be largely "attributed to the fact that newspapers are regarded primarily as conveyors of news through text. […] graphical news and data is subsidiary to the textual portion of a newspaper article because of reader, reporter, and editor expectations that a thorough and valid newspaper report ought to be primarily textual and attributed to a writer rather than a graphic artist." Yet she claims that people's expectations differ on the web. Indeed, many web-based infographics and visualizations now exist independently from other media—as attested in various online visualization galleries and community websites (*e.g.*, visualisingdata.com, visualcomplexity.com, or visualizing.org), and thanks to the recent development of specific and easily accessible toolkits and programming libraries (*e.g.*, [[47], [60], [100]]). However, it seems that in many cases still, the most compelling mass-reaching uses of infographics and visualization are dependent on other means of communication. For example, the effectiveness of Al Gore's evolution of CO2 emissions graph, presented in his *Inconvenient Truth* [59], is largely dependent on his staging of the magnitude of the numbers; he uses a platform elevator to reach up to the highest point in the graph. Similarly, Hans Rosling's energetic TED Talks [106] are surely the most engaging way to discover his *GapMinder* visualizations [86].

Finally, beyond traditional mass-media graphics, many visual representations of data now also reach out to other contexts and environments; Pousman & Stasko have proposed a categorization of these visualizations, which they qualify as *casual information visualizations* [226]. These can be found in museums, on the web, or in people's homes; they generally help reflect on everyday life, and sometimes even challenge the core notion of infovis as an amplifier of cognition.

2.3.2 Infovis for *the* People

Information visualization for *the* people [147], or for *the masses* [161], is a relatively new research topic in the infovis community. The term has been proposed to define the outreach of information visualization to casual audiences. In his capstone presentation at the 2007 IEEE Infovis conference, Stephen Few stated that "there are a lot of people in the world out there who stand to benefit from what we [the Infovis community] do, and the view that they have of infovis is usually different from our own. It's important that we understand it if we wish to make our efforts count in the world" [156]. Viégas *et al.* have also advocated for the democratization of information visualization [267], by both providing technology to the broadest possible audience, and encouraging a "democratic deliberative style of data analysis."

The difference between information visualization for *the* people and the other mass-reaching graphics mentioned above is essentially the potential that the first provides for interaction and exploration. Danziger proposes the following definition: *"information visualization for the people is not hardcore analytics, nor information art, nor any other perspective on data representation that doesn't facilitate [information interaction]. It is user-centered information visualization designed in a way that the average user can both do relevant analysis and enjoy the experience of interacting with information"* [147].

Danziger also identifies the four following design dimensions for infovis for *the* people:

* Dl—*Semantic design:* infovis for *the* people should be designed as a visual language based on a sign system, which people should be able to interpret or "read";
* Da—*Aesthetic and affective design:* infovis for *the* people should make use of 'artistic' or affective design principles, to evoke emotional engagement with its content;
* Dn—*Narrative design:* infovis for *the* people should be designed to convey information in the form of a story; and
* Ds—*Social design:* infovis for *the* people should incorporate social media features to facilitate collaborative analysis and the emergence of collective intelligence, while "humanizing the process of interacting with data".

Similarly, Murray proposes three *avenues of engagement* for information visualization design, which aim to bring people into the data, and to communicate their meaning through stories; these avenues are: *aesthetics,* *narrative,* and *interaction* [105]. Needless to say, aesthetics correspond to Da, and narrative to Dn, while interaction enables people to explore the data on their own. Note that the interaction avenue is implied in Danziger's definition of infovis for *the* people [147], which is why I will stick with his dimensions throughout the rest of this section. In the following paragraphs, I relate these to the different sub-costs of Ce.

Dl is first related to the **literacy cost**. While Danziger mentions that the design of visualizations as sign systems has been heavily studied, referring mainly to the works of Bertin [118], Card *et al.* [131], and Mackinlay [199] (see Section 1.1.2), and while research in graph comprehension has uncovered the higher-level cognitive processes behind the reading of graphs (see Section 2.2.1), I argue that the problem of *visualization literacy* has still been under-explored. If people are not used to interpreting visual mappings as properties of the underlying data, they will likely not engage with information visualizations. In his keynote presentation at the 2007 IEEE Infovis conference, New York Times deputy graphics director Matt Ericsson evoked this issue, stating that many readers have trouble interpreting scatterplots when the X axis does not encode time (in [147]).

Dl is also related to the **context-interpretation cost**. As discussed in Section 2.2.2, visualizations are often designed to be content-agnostic, in order to fit as many application cases as possible. While this abstraction may encourage user-exploration in some cases, Danziger points out that there is "a fine line between visualizations that are mysterious and intriguing, and ones that are mysterious and incomprehensible" [147]. He then goes on to say that "unfortunately, much of current infovis falls in the latter category with respect to general audiences, either due to excessive abstraction, excessive complexity, or simply lack of explanatory documentation." Similarly, Murray argues that "an honest representation [of data—in Tufte's sense] can be so abstract and so reductive as to be inaccessible to viewers" [105].

Da is not directly related to the sub-costs of Ce. Addressing the **context-interpretation cost** may increase the attractiveness of a visualization (see Section 2.2.2), but I believe aesthetics are a more general issue.

O'Brien & Toms have found that aesthetics predict perceived usability [214], and Tractinsky *et al.* claim that "What is beautiful is usable" [251]; and, as mentioned in Section 2.1.2, a visualization should first of all be aesthetically appealing for people to engage with it.

Finally, while Dn and Ds are not directly related to the sub-costs of Ce either, I hypothesize they may be useful for addressing the **initial incentive cost**. They may also help overcome the **context-interpretation cost** and the **perceived interactivity cost**, as these design dimensions can be put to use for illustrating the topic of a dataset, and for creating more or less explicit tutorials, which can guide users though the different interactive features of a visualization.

In the following subsections, I detail previous research conducted along these different design dimensions, and shed more light on how they relate to the sub-costs of Ce.

2.3.2.1 *Semantic Design*

Semantic design is about helping viewers interpret a visualization; it is an important dimension of *ambient information visualization* design (or *ambient information system/display* design). Mankoff *et al.* define such ambient displays as *"abstract and aesthetic peripheral display portraying non-critical information on the periphery of a user's attention"* [201]. In an attempt to evaluate the comprehension viewers have of such displays, Skog *et al.* have proposed a three-step scale [244] in which a viewer should realize: 1) *that* something is visualized (*e.g.*, "Does the viewer know that the display is an information visualization and not simply decoration?"); 2) *what* is visualized (*e.g.*, "Can the viewer tell what data the visualization reflects?"); and 3) *how* the data are visualized (*e.g.*, "Can the viewer read and interpret the visualization correctly?").

Understanding that something is visualized, and how the data are visualized is related to the **literacy cost**. While Skog *et al.*'s scale directly concerns ambient visualizations, which are generally less focused on presenting data for analysis than other information visualization systems for *the* people, I argue casual audiences may not always overcome these steps, even with simpler/more traditional graphics. Some people may be used to seeing business graphics and may realize that they present some sort of data, but they may not be used to trying to understand how the data is mapped into visual form, or, for that reason, what it means. As an analogy, someone illiterate may realize that a block of text presents some sort of information (due to general cultural understanding), but does not know how to parse the representation. Interestingly however, although true conclusions are hard to formulate based on their qualitative study which only included six participants, Skog *et al.* mention that the only person who did not realize that data was visualized by the ambient display they were evaluating was unfamiliar with the dataset itself (a bus-line departure times); all others had prior experience with and/or knowledge of the data [244]. This suggests that having background knowledge of the data can help people identify that something is being visualized, and it may even help them identify how it is visualized.

Understanding what is visualized is related to the **context-interpretation cost**. Here too, I argue this issue goes beyond the spectrum of ambient visualizations. To address it, Pousman and Stasko [225] identify a *representational fidelity* dimension in ambient system design, which they describe as the level at which patterns, pictures, words, or sounds produced by the system *stand* for the things they represent, *i.e.,* the data. They use the language of semiotics (see Section 2.2.2) to analyze this, and propose the five following levels of fidelity, ranging from high to low: 1) *indexical* (*e.g.,* measuring instruments, maps, or photographs); 2) *iconic-1* (*e.g.,* drawings, doodles, or caricatures); 3) *iconic-2* (*e.g.,* metaphors); 4) *symbolic-1* (*e.g.,* language symbols like letters and numbers); and 5) *symbolic-2* (*e.g.,* abstract symbols).

Outside the subfield of ambient visualizations, the work of Otto Neurath is a common reference for semantic design [[147], [180]]; already in the 1930's, his *International System Of TYpographic Picture Education* (ISOTYPE) aimed to improve cultural communication, and to democratize cultural life beyond the traditional limitations of cultural, social, or educational backgrounds [210]. This system essentially consists in assembling easily recognizable or interpretable pictograms (a combination of icons, indexes, and symbols—see Section 2.2.2) to communicate statistical values. Huron *et al.* have decomposed his assembly process into: 1) showing numeracy *via* countable units (which they name *tokens*); 2) using pictograms to encode each of these units (the *token grammar*); and 3) composing these pictograms into comparable layouts (the *assembly model*) [180]. They have advocated for employing a *constructive* approach to visualization design, based on the use of such tokens.

Finally, somewhat related to Huron *et al.*'s approach is Chevalier *et al.*'s *Concrete Scales* framework [135]. Inspired by the work of many independent designers, they propose that using visual metaphors to which viewers can relate helps communicate complex measures. For example, showing the amount of sugar in an orange is more easily understood when the graphic presents an orange next to a stack of sugar cubes. This approach is very literal, but it has the advantage of exposing upfront what the data are about, *i.e.*, oranges and sugar.

2.3.2.2 *Aesthetic Design*

Aesthetic design is about generating an immediate positive emotional response in viewers. According to Mankoff *et al.*'s definition [201] (see Section 2.3.2.1), aesthetics are another important dimension of ambient visualization design [[193], [225], [246]]. Pousman & Stasko suggest that while aesthetic appeal is a purely subjective experience, it can be identified as a designer's *intention* to produce an "art worthy of contemplation" [225].

In line with this, Wood *et al.* hypothesize that the sketchy rendering of a visualization can be aesthetically appealing, as it conveys a notion of intension—the appearance suggests manual effort, and therefore that the graphic was produced "for a purpose" [275]. In relation to Tversky's work [256], they also mention that sketches generally omit irrelevant information, which simplifies the interpretation process and reveals the sketcher's conception of the domain. Finally, they suggest that sketchiness reinforces the perception of simplicity, which may reduce the expectation of cognitive effort for interpreting a visualization. Viégas & Wattenberg also highlight this idea of intention [266], and claim that artistic visualizations are more about an artist's intent than the actual "surface aesthetics".

From a broader perspective, Lau & Vande Moere propose a model for *information aesthetics* [190], at the intersection between information visualization and visualization art. The model's main dimensions are *data focus* and *mapping technique;* the first is a spectrum between intrinsic and extrinsic focus, and the second another spectrum between direct and interpretative techniques. *Intrinsic data focus* aims to facilitate insight into data by using cognitively effective visual mappings, and allows viewers to discover useful patterns in those data. *Extrinsic data focus* aims to facilitate communication of meaning related to, or underlying data, and encourages viewers to develop personal interpretations and reflections. *Direct mapping techniques* aim for the most ideal representation for a given data type; they are inversible, meaning viewers can infer the underlying data values from the visual representation. Finally, *interpretive mapping techniques* involve more subjective design decisions and stylistic influences; they cannot be inversed, meaning that viewers will have more trouble comprehending the underlying data values or patterns. Lau & Vande Moere show that data focus and mapping technique are qualitatively correlated in many visualization systems or designs, meaning the data focus generally determines the mapping technique, and *vice versa;* they also identify how several subfields of information aesthetics relate to this correlation.

Although not directly related to information visualization, Jennings also proposes a *prescriptive aesthetic* framework [184] for engaging and immersing users in websites. This framework combines characteristics of *aesthetic experience* [143] with design recommendations collected through a survey conducted with "leaders in the industry" of educational game environments. The main characteristics of the framework are: unity, attention or object directedness, active discovery, affect, and intrinsic gratification. *Unity* concerns providing the user with a holistic environment, in which s/he can experience all the other characteristics. *Attention of object directed-*

ness concerns elements that bring about focus, or a desire to proceed with an activity. *Active discovery* concerns challenging the user to make sense of potentially conflicting stimuli, in order to improve his/her skills and knowledge. *Affect* concerns the emotional involvement a user makes, in order to be immersed in an environment, and to sustain this immersion. Finally, *intrinsic gratification* concerns the feeling of pleasure a user gains from conducting an activity, where the reward becomes the activity itself (see Flow theory—Section 2.2.4). Jennings' full framework is provided in Appendix D. Interestingly, it extends the single aesthetics perspective in several ways. For example, the *Familiarity* dimension of the focused attention and object directedness characteristic stresses that users pay more attention to what they know and understand. This can be related to Skog *et al.*'s finding that the only person who did realize that data was visualized was not familiar with the dataset [244] (see Section 2.3.2.1). I relate this to the **context-interpretation cost**, and I hypothesize that helping a user relate to a dataset is an important attribute of semantic design (not of aesthetic design). In addition, the *Personal Motivation* dimension of the intrinsic gratification characteristic emphasizes the fact that users often come to a website with personal motivation, but that the website should include innovative techniques to sustain this motivation. While I speak of the **initial incentive cost**, these dimensions are related in the sense that I consider that information seekers come across an information visualization with the motivation of finding new information (or of simply viewing the visualization), but may need extra 'initial' incentives for actively exploring the data–this can be considered as a way to 'sustain' the original motivation of finding new information.

Finally, Byron & Wattenberg raise the concern of a trade-off between aesthetics and utilitarian consideration, as aesthetics may sometimes impede legibility [128]. Their conclusion is that the balance should be made according to contextual needs. In cases where the audience is fixed and captive, there may be no need to compromise legibility. However, in other cases, it may be worth while to prioritize aesthetics to broaden the appeal of a graphic before people engage in exploring the data.

2.3.2.3 *Narrative Design*

Narrative design is about building a (visual) narrative around data. Murray mentions increasing popularity of a format where authors (or designers) tell a story first, before letting users dive into more detail if they are interested [105]. This suggests that narrative design can be used to build interest in users, thus helping them overcome the **initial incentive cost**. In addition, narrative design can be used to immerse users in the 'story' of the data, thus helping them overcome the **context-interpretation cost**. Such *narrative visualization* [[179], [238]] formats have encouraged a new line of research in infovis, which explores the potential for storytelling with information visualizations.

Hullman & Diakoplous define narrative information visualizations as *"a style of visualization that often explores the interplay between aspects of both explorative and communicative visualization. They typically rely on a combination of persuasive, rhetorical techniques to convey an intended story to users as well as exploratory, dialectic strategies aimed at providing the user with control over the insights she gains from interaction"* [179]. This interplay raises a tension previously identified by Segel & Heer, between author-driven and reader-driven scenarios [238]. *Author-driven* scenarios follow a linear structure intended by the author; in their most extreme incarnation, they provide no interaction. On the contrary, *reader-driven* scenarios give control to the person receiving the information by providing an open system, and allowing free interaction. Note that interactive narrative visualizations rarely fall directly into either of these categories, but rather somewhere along a spectrum between the two.

New York Times graphic editors Mike Bostock and Shan Carter make a similar distinction between explanatory and exploratory graphics in [62]. They describe *explanatory graphics* as having the advantage of exposing up-front what the main insights from the data are, without making people have to "work" for them; while, *exploratory graphics* require some amount of effort to extract interesting information. Bostock & Carter suggest that the success of a journalistic exploratory graphic largely depends on contextual factors. For example, Carter mentions that elections are a "unique" opportunity for such graphics, because people are inherently interested in the topic; they have their own "hunches" about the data, which encourages them to explore the visualization. This suggests that such contextual factors can help overcome the **initial incentive cost**. However, even in such contexts, Bostock reports possible issues with the **perceived interactivity cost**, as several viewers of the *512 Paths to the White House* graphic [92] did not realize the visualization was interactive [62].

Segel & Heer also propose a design space for narrative elements [[238], Fig.7], and identify three common structures of interactive narrative visualizations: the Martini Glass structure, the Interactive Slideshow, and the Drill-Down story. The *Martini Glass* has a two-stage structure: first, the user goes through a relatively heavily author-driven scenario, in which the visualization is introduced through the use of text, annotations, nicely crafted animations, or interesting and evocative views. Second, when the author's intended narrative is complete, the user is put in charge and can actively explore the visualization following whichever path s/he considers most interesting. Thus, the authoring segment should function as a "jumping off point for the reader's interaction" [238], *i.e.,* as a way to overcome the **initial incentive cost**. The *Interactive Slideshow* structure follows a standard slideshow format, and allows for mid-narrative interaction within each slide. These may follow the Martini Glass structure by presenting the author's intended story before inviting the user to interact with the display. Thus, this structure is more balanced between the author- and reader-driven approaches. The *Drill-Down* story presents a general theme, and lets users select among particular instances of that theme to reveal details and backstories. While this structure emphasizes reader-driven scenarios, it also requires a significant amount of authoring to determine the possible interactions, the stories to reveal, and the details to include in each story. Finally, Segel & Heer also mention that narrative visualization can include "tacit tutorials" of available interactions, which can compensate for the **perceived interactivity cost**.

Hullman & Diakopoulos extend the notion of author-driven scenarios, and propose a framework for *visualization rhetoric* [179]. They characterize a certain number of design choices a visualization author can make, in order to create what Narratologists refer to as the narrator's *voice* [182](2). They call these design choices the *framing effect* in narrative visualizations, and identify the following dimensions: information access rhetoric, provenance rhetoric, mapping rhetoric, linguistic-based rhetoric, and procedural rhetoric. *Information access rhetoric* is related to the designer's choice of what data to present. *Provenance rhetoric* is related to a designer's will for transparency and ethics—I would also add the fact that providing provenance can help as a case-defender; if data come from a well-acknowledged and trusted provider, viewers are more likely to 'believe' the information than if data come from an unknown provider. *Mapping rhetoric* is related to the designer's choice of visual mappings for the data. *Linguistic-based rhetoric* is related the stylistic techniques implemented in the textual layers of a visualization. Finally, *procedural rhetoric* is related to the way in the which the narrative unfolds—this can be related to Segel & Heer's common structures of interactive narrative visualizations [238] (mentioned above). Interestingly, Wood *et al.* relate their hypothesis of designer's intentions being a conduit for aesthetic appeal to this framing effect [275] (see Section 2.3.2.2).

Finally, as an offspring of narrative visualizations, Diakopoulous *et al.* have developed the concept of *game-y graphics* [[149], [150]]. They explore the intersection between visual analytics and game design, and propose that including game mechanics to visualizations can help motivate users explore data, by biasing their attention and interaction *via* the goals and representations embedded in the game mechanics. This suggests that gaming can also be an interesting way of overcoming the **initial incentive cost**, as it provides an external motivation or goal (*i.e.,* to complete the game or to compete with others), which may then be internalized by players.

2 Note that the narrator's voice can be that of the author her/himself, or that of a fictional character, which can be set within or outside the story.

2.3.2.4 Social Design

Social design is about enabling social interactions around visualizations. Wattenberg has proposed the term *social data-analysis* to define such discussions and social exchanges [269]. He claims that exploring a dataset can become a social activity, and he identifies several different roles people may assume [270]. Interestingly, he hypothesizes that viewing exploratory data analysis as a social activity can help engage people who are not necessarily intrinsically interested in the data. This can be related to Ryan & Deci's suggestion of providing people with a sense of relatedness to help them internalize their motivations [235] (see Section 2.2.4), and indicates that social design may also help overcome the **initial incentive cost**. In addition, it may enable collective sense-making processes, which may sustain people's engagement.

Wattenberg proposes the following perspectives for designing visualizations that support social data-analysis [269]—these should:

* provide means for establishing *common ground* and *unique perspectives:* a visualization should establish common ground through the type of data it presents and/or the way it presents them; this helps people relate to the topic. However, the visualization should also allow people to find and share unique perspectives.
* provide an *expressive spectator interface:* in the event where several people use a visualization together, this visualization should be expressive, so that the people who are not directly interacting with it, *i.e.*, spectators, can understand what is going on, and can make suggestions to the person who is interacting with the display.
* enable *discovery transfer:* a visualization should allow people to export and/or point to a specific view to help communicate insights and findings.

Interestingly, the common ground and unique perspectives principle resembles Carter's description of a "unique" opportunity for exploratory graphics [62] (see Section 2.3.2.3), where people are inherently interested in the topic and share the same general understanding of it (the common ground), yet have their own personal "hunches" about the data (the unique perspective), which they may be willing to share.

Heer adds several other considerations [172]: *connecting to data, i.e.,* enabling users to establish a personal connection with data; *conversation and community formation, i.e.,* providing a space in which communication can occur; and *exploring boundaries, i.e.,* understanding the different ways in which users may engage with a visualization, other than in the ones intended by the designer. He emphasizes that visualizations should not only be considered as external cognitive artifacts, but also as social artifacts.

Based on these considerations, Viégas & Wattenberg have encouraged the development of a new line of research on *Communication-Minded Visualization* (CMV) [265]. Taking inspirations from the *Computer Supported Cooperative Work* (CSCW) community, they propose a framework for visualization design that supports collaboration between users, based on the location of analysts and on the time of their analysis. They stress that there has been very little research in *distant-asynchronous* collaborations using visualizations—a dimension they call *asynchronous sharing.* They also mention that providing means for establishing common ground and *deixis* are crucial to successful user-collaborations.

Willett *et al.* have taken several steps in the direction of CMV. First, the *CommentSpace* component for collaborative visual analytics [271] enables analysts to structure discussions around visualizations using tags and links to annotate comments. Second, Willett *et al.* have proposed seven strategies for optimizing crowdsourced data-analysis [272], based on a set of five problems that can reduce the quality of crowdsourced explanations. Interestingly, their *unclear expectations* problem (problem 2) seems to highlight the **literacy cost**. Third and last, Willet *et al.* have formulated strategies for identifying redundancy and showing provenance of explanations in crowdsourced data-analysis [273], and have developed an *explanation-management* interface, with which analysts can quickly group and filter crowdsourced explanations based on plausibility. While this work is interesting, it inherently assumes a hierarchical structure where analysts (at the top) provide workers (at the bottom) with a series of specific analytic tasks to perform. In this dissertation, I assume a more ‘grassroots’ approach in which people should engage at their own pace in the exploration of data.

2.3.2.5 *Summary*

Overall, these different design dimensions can be put to good use for helping casual audiences overcome the sub-costs of Ce. They also highlight the intricate relatedness of these sub-costs. Semantic design can be used to address both the **literacy cost** and the **context-interpretation cost**; aesthetic design is a general concern that is indirectly associated with all the sub-costs; narrative design can be used to address the **context interpretation cost**, **perceived interactivity cost**, and **initial incentive cost**; and social design can also be used to address the **initial incentive cost**. Moreover, these dimensions are themselves closely connected. Semantic design can help create an aesthetic experience, which may assist aesthetic design. Likewise, narrative design can use elements of semantic design to create a compelling story, which can then also help create an aesthetic experience. Finally, social design can assist narrative design if the ‘story’ is built with the different social exchanges that occur around a visualization.

2.4 Success Stories and Acknowledged Failures

Having described previous work conducted along the four design dimensions of infovis for *the* people, and having discussed how these may address the sub-costs of Ce, I now turn to a review of several online information visualizations, which were inspired by (or that have inspired) these design dimensions. In this section, I present a series of popular success stories and acknowledged failures, concentrating on visualizations that have the potential for data exploration, *i.e.*, not on communicative graphics. While it is difficult to separate out all the reasons for success, and to speculate on the effect of the sub-costs of Ce that may not have been addressed by successful visualizations (since academic papers generally tend to minimize—if not to completely leave out—possible failure indicators), it is interesting to consider success through the sub-costs such visualizations do address. In addition, identifying the sub-costs that may have led to failure can highlight which of these are more critical, or difficult to address.

2.4.1 The Name Voyager

A first popular success story is Wattenberg's *Name Voyager* [76]. This visualization shows the evolution of baby name popularities since 1900 in the form of a stacked graph; it uses three simple widgets for filtering the data, namely a prominent search bar that reacts directly to keystrokes and updates the chart in a nicely animated way, and three radio buttons to filter out genders. The visualization website was visited more than 500,000 times in the two first weeks after it was launched, and maintained an average 10,000 visits a day after the two first moths.

While the visualization does not include any social design features *per se* (as presented in Section 2.3.2.4), one of the main attributes of its success is the social activity it generated. Many people have discussed the visualization in forums and blogs, while pointing back to the website; some have even set pattern-finding challenges [269]. This undoubtably provided new users with initial motivations for exploration, and helped them overcome the **initial incentive cost**. In addition, the prominent search bar addresses the **perceived interactivity cost**, and the simple cultural color scheme used (*i.e.*, blue for boy names and pink for girl names) addresses to some extent the **context-interpretation cost**. This is interesting, as it suggests that free visual variables may be used to help interpret the data, making it possible to avoid adding chart junk to the visual representation.

2.4.2 The GapMinder

Hans Rosling's *Gapminder* is another popular visualization website [86] that provides a main visualization component (the *Gapminder World*), as well as several other narrative visualizations that are available for download. It is a "modern 'museum' that helps making the world understandable, using the Internet" [86]. However, while the *Gapminder* has been made famous by Hans Rosling's popular TED talks [106], the online visualizations themselves are not mentioned very often.

The success of the *Gapminder* can be heavily attributed to Hans Rosling's effective use of storytelling, which has been transposed to several of the downloadable narrative visualizations (*e.g.*, the *Human Development Trends, 2005* and the *Has the World Become a Better Place?* visualizations). These generally use an interactive slideshow format (see Section 2.3.2.3), which guides the user step by step through the different visual features of the display, thereby addressing both the **literacy cost** and the **context-interpretation cost**. The visualizations also provide interesting cues to help with the **perceived interactivity cost**, like a 'fake' pointer which animates in and out of view to suggest where the user might click.

However, these attributes are not implemented in the main visualization component, which makes me believe that the pedagogical approach used in the narrative visualizations can serve as an entry point by providing users with initial incentives to go explore the *Gapminder World.*

2.4.3 Sense.us

Heer *et al.*'s *Sense.us* was yet another popular visualization website—although it is no longer online—that provided a set of visualizations of 150 years worth of United States census data, which users where invited to explore, comment, and annotate. The main novelty, which was "probably the most effective and well-liked novel feature", was a *doubly linked* discussion model [173]. This model enabled users to lead 'independent' discussions in a standard forum-like interface, while providing them with means to annotate the visualization, and to link views with their comments. Thus, when a user decided to inspect a comment, the appropriate view of the data was displayed; and *vice-versa,* when the user explored the visualization and came across a certain parametrization that had been discussed, the comments would appear in the forum-like interface.

I hypothesize that the success of *Sense.us* may be attributed to the fact that social interactions can help overcome the **initial incentive cost**. In addition, the *doubly-linked* discussion model may have provided a sense of unity (see Appendix D), which enhanced users' aesthetic experience.

2.4.4 Many Eyes

IBM's *Many Eyes* is a famous visualization platform(3) that allows users to upload and transform their data to fit analytic needs; to choose from several visualizations to encode the data; and to publish their final graphic. The goal is to enable discussions, both about data and visualizations.

Many Eyes is very different from the visualization websites presented above, as it requires consideration of all of van Wijk's costs [263] (see Section 1.1.2.2). However, I will remain focused on Ce. I consider the platform from an information seeker's point of view who is simply browsing through the different visualizations created by others—even though it

3 I use the term "platform" here because Many Eyes is more than a 'simple' visualization website: it provides tools to create and publish visualizations within the platform.

may be argued that the purpose of *Many Eyes* is to create a community around data and information visualization, rather than just providing visualizations that any 'outsider' can explore. As such, my review does not concern the platform itself, but rather the different visualizations it hosts.

From this perspective, a first striking observation is that it is extremely difficult to understand what most visualizations are about. Many use the same encodings, and look alike. This means that the **context-interpretation cost** is very high. Furthermore, most visualizations do not include much more than a title to describe what the data are about, and these titles are often quite obscure. For example, a treemap found on the platform is entitled "Student Feedback Survey Summary for CSE at UTA in 2014 Fall", and its description is "Student feedback survey summary" [73]. Who, other than the author, is meant to understand this? In addition, while many of the visual representations used are standard business graphics, like bar charts, pie charts, and line graphs, some are more advanced, like treemaps, scatterplots, boxplots, or starplots. This requires viewers to be well accustomed with visualizations, in order to switch from one type of representation to another, and means that the **literacy cost** is likely to be high. Also, in some cases the snapshot shown on the main page uses a different kind of visual representation than the actual visualization (*e.g.,* [73]).

Furthermore, while several interaction techniques are automatically integrated to each visualization, it is difficult to know which ones can be performed, when they can be performed, and if they are useful. Typically, drop down menus are provided beneath each visualization to help viewers modify the encoding parameters, but these often offer only a single dimension to chose from (the one already displayed). Similarly, all visualizations allow for drag-selections, but in several cases this is useless. Finally, in some cases the user can interact with the legend of the graphic (*e.g.,* in [72]), which is not a common feature, and which needs to be 'discovered.' Overall, these issues may be confusing, and may lead users to disregard interactivity even when it is useful. Thus, the **perceived interactivity cost** is also high.

Finally, many visualizations are not commented, and do not provide any initial incentives for exploring the data they present. Admittedly, this may be due to the fact that at the time I conducted this review of the platform, it was being updated to a new system and not all the content was available. In addition, a lot of the data visualized on *Many Eyes* only concerns small groups of people, which makes it difficult for outsiders to find motivation to explore the datasets. Thus, at least in this state, the **initial incentive cost** is also high.

2.4.5 Verifiable.com

Verifiable was a visualization platform that ran between 2007 and 2010; like *Many Eyes,* it allowed users to upload data and to create their own visualizations, but in a more "analytical perspective" [37]. In a blog-post on his personal website, Ex-President Stuart Roseman describes the reasons why he believes the project failed [13]. Among other technical issues, he mentions that people are only interested in what they can immediately relate to. In a later blog-post, he also considers that the timing of *Verifiable* "might have been a little early," since "everyone is now talking about big data, big data, big data" [61].

Kosara discusses the issues raised by the *Verifiable* case [37], and mentions that "the idea of visualization for the masses is a good one, but not if it also done by the masses." He argues that it is important to make sure people actually have an idea of what to do with visualizations, before attempting to assist them in their design. This is directly related to the subcosts of Ce, and although Kosara's thoughts were written almost five years ago, I posit they are still valid.

2.4.6 Swivel.com

Like *Verifiable*, *Swivel* was an online visualization platform that ran between 2006 and 2010; it was referred to by bloggers as the "Youtube for data." In an interview with Robert Kosara, co-founders Brian Mulloy and Dimitry Dimov discuss their initial motivations, and the reasons why they shut down the platform [39]. Their original hypothesis was that "there are a whole bunch of people who are not analysts or statisticians, or visualization experts, who would really benefit from seeing, and engaging with, statistics. And if we made it engaging, they would." However, very few people did engage with the platform—less than ten customers payed for the service. Backing-up Roseman's description of the problems that lead to the end of *Verifiable*, Mulloy attributes this to the overall lack of interest "people who are not inherently biased towards working with datasets" have in exploring data. Like in the *Verifiable* case, this suggests that the sub-costs of Ce must be addressed before information visualization can truly be democratized.

2.5 Conclusion

In this chapter, I have introduced the concept of engagement, and have shown how the sub-costs of Ce may articulated according to a four-stage-model of engagement (see FIGURE 2.2). I have also presented how the **literacy cost** relates to graph comprehension, how the **context-interpretation cost** can be considered through the language of semiotics, how the **perceived interactivity cost** relates to the concept of perceived affordances, and how the **initial incentive cost** relates to motivation.

While I have purposefully distinguished each of the sub-costs of Ce, it is interesting to note how this theoretical understanding of the constructs behind them highlights possible relationships. The **context-interpretation cost** may be connected to the **perceived interactivity cost**, as addressing the first can increase the attractiveness of a visualization, *i.e.*, its aesthetic appeal (see Section 2.2.2), predicting its perceived usability (see Section 2.1.1); this may then entice users to 'try out' the visualization, which may compensate for the second cost. Likewise, the **literacy cost** and the **context-interpretation cost** may be connected to the **initial incentive cost**, as the low-level question posing process required to parse a visual representation, *e.g.*, at the elementary level described by Bertin (see Section 2.2.1), may lead to higher-level question posing processes, which may then compensate for the latter cost.

In addition, I have presented previous work on information visualization for *the* people, which I have also related to the sub-costs of Ce; and I have reviewed several success stories and acknowledged failures of visualization websites for *the* people. This has provided insight into how important it is to consider these different costs when designing visualizations for casual audiences.

While I do not argue with the successful nature of some of the visualizations reviewed in Section 2.4, it is important to note that no published work accounts for what users' actually *do* with these visualizations, *i.e.*, how deeply they explore the data, if they conduct their exploration in a single session or in several, *etc.* None report tracking behavioral data on a large scale. Although analyzing comments posted inside or outside a visualization website can provide insight into what users understand about the data, these indicators are often 'noisy' and limited. People tend not to venture into technical details in their comments, and may completely leave out issues that they did not even realize were there (as reported in [62]). This lack of lower-level behavioral information makes it impossible to know whether people actually *use* visualizations to gain insight. In Chapter 6 of this dissertation, I attempt to operationalize the measure of engagement using such data, and I establish several metrics that can serve as proxies of user-engagement.

Finally, while in this chapter I have discussed the relatedness of several of the sub-costs of Ce, both in theory and through design solution that may address them, and while proper consideration of some sub-costs may compensate for the absence of others, I continue to treat them individually throughout this dissertation, as I a believe it is important to find ways to address each of them before attempting to combine them in different ways.

Chapter 3

Measuring the Literacy cost: A Principled Way of Assessing Visualization Literacy

In April 2012, Jason Oberholtzer posted an article describing two charts that portray Portuguese historical, political, and economic data [14]. While acknowledging that he is not an expert on those topics, Oberholtzer claims that thanks to the charts, he feels like he has "a well-founded opinion on the country." He attributes this to the simplicity and efficacy of the charts. He then concludes by stating: "Here's the beauty of charts. We all get it, right?" But do we *all* really get it? Although the number of people familiar with visualization continues to grow, it is still difficult to estimate anyone's ability to read graphs and charts; and as I have discussed in the previous chapters, this **literacy cost** can be an important drawback for casual audiences.

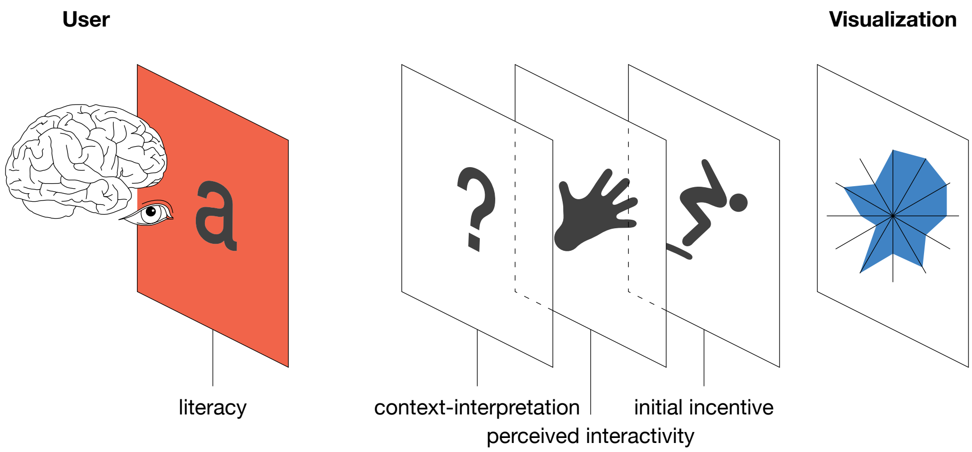


FIGURE 3.1: The **literacy cost**.

This chapter is based on a published paper entitled *A Principled Way of Assessing Visualization Literacy* [122], so any use of the term "we" refers to myself, Ronald A. Rensink, Enrico Bertini, and Jean-Daniel Fekete. It focuses on the **literacy cost** by addressing the following research question:

Q1: How can a designer know the level of understanding an audience has of different visual representations of data?

To answer this, we present the design and calibration of a set of *visualization literacy* (VL) tests for line graphs, bar charts, and scatterplots. To generate these tests, we propose a method based on *Item Response Theory* (IRT). Traditionally, IRT has been used to assess examinees' abilities *via* predefined tests and surveys in areas such as education [187], social sciences [159], and medicine [43]. Our method uses IRT in two ways: first, in a *design phase,* we evaluate the relevance of potential test items; and second, in an *assessment phase,* we measure users' abilities to extract information from graphical representations. Based on these measures, we then propose a series of tests for fast, effective, and scalable web-based use. The great benefit of this method is that it inherits IRT's property of making ability assessments that are based not only on raw scores, but on a model that captures where a user is situated along a continuous *latent trait* scale (*e.g.*, along a scale ranging from inability to high ability to use various graphical representations).

As such, the main contributions of this chapter are as follows:

* a practical definition of visualization literacy;
* a method for: 1) assessing the relevance of visualization literacy test items, 2) assessing an examinee's level of visualization literacy, 3) creating fast and effective assessments of visualization literacy for well established visualization techniques and tasks; and
* an implementation of four online tests, based on our method.

Note that we adopt an evaluation approach here, as our immediate motivation for this work is to design a series of tests that can help infovis researchers detect low-ability participants when conducting online studies, in order to avoid possible confounds in their data. This requires the tests to be short, reliable, and easy to administer. However, such tests can also be applied to many other situations, such as:

* designers who want to know how capable of understanding visualizations their targeted audience is;
* teachers who want to make an assessment of the acquired knowledge of freshmen;
* practitioners who need to hire capable analysts; and
* education policy-makers who may want to set a standard for visualization literacy.

This chapter is organized in the following way. It begins with a background section that extends Section 2.2.1; it defines the concept of *literacy* and discusses some of its best-known forms. Also introduced are the concepts behind Item Response Theory. Section 3.2 then presents the basic elements of our approach. Section 3.3 shows how these can be used to create and administer two visualization literacy tests using line graphs. In Section 3.4, our method is extended to bar charts and scatterplots. Section 3.5 describes how our method can be used to redesign fast, effective, and scalable web-based tests. Finally, Section 3.6 provides a set of take-away guidelines for the development of future tests.

3.1 Background

Very few studies investigate the ability of a user to extract information from a graphical representation such as a line graph or a bar chart. And of those that do, most make only higher-level assessments: they use such representations as a way to test mathematical skills, or the ability to handle uncertainty [[6], [22], [25], [89], [216]]. A few attempts do focus more on the interpretation of graphically-represented quantities [[58], [94]], but they base their assessments only on raw scores and limited test items. This makes it difficult to create a true measure of visualization literacy.

3.1.1 Literacy

3.1.1.1 *Definition*

The online Oxford dictionary defines literacy as *"the ability to read and write."* While historically this term has been closely tied to its textual dimension, it has grown to become a broader concept. Taylor proposes the following: *"Literacy is a gateway skill that opens to the potential for new learning and understanding"* [249]. Given this broader understanding, other forms of literacy can be distinguished. For example, *numeracy* was coined to describe the skills needed for reasoning and applying simple numerical concepts. It was intended to "represent the mirror image of [textual] literacy" [141], p. 269]. Like [textual] literacy, numeracy is a gateway skill.

With the advent of the Information Age, several new forms of literacy have emerged. *Computer literacy* "refers to basic keyboard skills, plus a working knowledge of how computer systems operate and of the general ways in which computers can be used" [217]. *Information literacy* is defined as the ability to "recognize when information is needed," and "the ability to locate, evaluate, and use effectively the needed information" [74]. *Media literacy* commonly relates to the "ability to access, analyze, evaluate and create media in a variety of forms" [90].

3.1.1.2 *Assessment*

Several literacy tests are currently in common use. The two most important are the UNESCO's *Literacy Assessment and Monitoring Programme* (LAMP) [99], and the OECD's *Programme for International Student Assessment* (PISA) [216]. Other international assessments include the *Adult Literacy and Lifeskills Survey* (ALL) [55], the *International Adult Literacy Survey* (IALS) [186], and the *Miller Word Identification Assessment* (MWIA) [83].

Assessments are also made on more local scales like the US *National Assessment of Adult Literacy* (NAAL) [113], the UK's Department for Education *Numeracy Skills Tests* [22], or the University of Kent's *Numerical Reasoning Test* [89]. Most of these tests, however, take basic literacy skills for granted, and focus on higher-level assessments. For example the PISA test is designed for 15 year-olds who are finishing compulsory education. This implies that examinees should have already learned—and still remember—the basic skills required for reading and counting. It is only when examinees clearly fail these tests that certain measures are deployed to evaluate the lower-level skills.

NAAL provides a set of 2 complementary tests for examinees who fail the main textual literacy test [113]: the *Fluency Addition to NAAL* (FAN) and *The Adult Literacy Supplemental Assessment* (ALSA). These focus on adults' ability to read single words and small passages. Meanwhile, MWIA tests whole-word dyslexia. It has 2 levels, each of which contains 2 lists of words, one *Holistic* and one *Phonetic*, that examinees are asked to read aloud. Evaluation is based on time spent reading and number of words missed. Proficient readers should find such tests extremely easy, while low ability readers should find them more challenging.

3.1.2 Visualization Literacy

3.1.2.1 *Definition*

The view of literacy as a gateway skill can also be applied to the extraction and manipulation of information from graphical representations. In particular, it can be the basis for what we will refer to as visualization literacy: *the ability to confidently use a given data visualization to translate questions specified in the data domain into visual queries in the visual domain, as well as interpreting visual patterns in the visual domain as properties in the data domain.*

This definition is related to several others that have been proposed concerning visual messages. For example, a long-standing and often neglected concept is *visual literacy*. This has been defined as the *"ability to understand, interpret and evaluate visual messages"* [126]. Visual literacy is rooted in semiotics, *i.e.*, the study of signs and sign processes, which distinguishes it from visualization literacy. While it has probably been the most important form of literacy to date, it is nowadays frowned upon, and general literacy tests do not take it into account. Meanwhile, Taylor [249] has advocated for the study of *visual information literacy*, and Wainer for that of *graphicacy* [268]. Depending on the context, these terms either refer to one's ability to read charts and diagrams, or to the merging of visual and information literacy teaching [109]. Because of this ambiguity, we prefer the more general term "visualization literacy."

3.1.2.2 *Assessment*

Relatively little has been done on the assessment of literacy involving graphical representations. However, interesting work has been done on measuring the perceptual abilities of a user to extract information from these. For example, various studies have demonstrated that users can perceive slope, curvature, dimensionality, and continuity in line graphs (see [140]). Correll *et al.* [140] have also shown that users can make judgements about aggregate properties of data using these graphs. Scatterplots have also received some attention. For example, studies have examined the ability of a user to determine Pearson correlation r [[119], [137], [205], [224], [230]]. Several interesting results have been obtained, such as general tendency to underestimate correlation, especially in the range $.2 < |r| < .6$, and an almost complete failure to perceive correlation when $|r| < .2$.

Concerning the outright assessment of literacy, the only relevant research work we know of is Wainer's study on the difference in graphicacy levels between third-, fourth-, and fifth-grade children [268]. He presents the design of an 8-item test using several visualizations, including line graphs and bar charts. He then describes his use of Item Response Theory [277] to score the test results, and shows the effectiveness of this method for assessing abilities. His conclusion is that children reach "adult levels of graphicacy" as soon as the fourth-grade, leaving "little room for further improvement." However, it is unclear what these "adult levels" are. If we look at textual literacy, some children are more literate than certain adults. People may also forget these skills if they do not regularly practice. Thus, while very useful, we consider Wainer's work to be limited. What is needed is a way to assess adult levels of visualization literacy.

3.1.3 Item Response Theory and the Rasch Model

Consider what we would like in an effective visualization literacy test. To begin with, it should cover a certain range of abilities, each of which could be measured by specific scores. Imagine such a test has 10 items, which are marked 1 when answered correctly, and 0 otherwise. Rob takes the test and gets a score of 2. Jenny also takes the test, and gets a score of 7. We would hope that this means that Jenny is better than Rob at reading graphs. In addition, we would expect that if Rob and Jenny were to take the test again, both would get approximately the same scores, or at least that Jenny would still get a higher score than Rob. We would also expect that whatever visualization literacy test Rob and Jenny both take, Jenny will always be better than Rob.

Now imagine that Chris takes the test and also gets a score of 2. If we based our judgement on this *raw* score, we would assume that Chris is as bad as Rob at reading graphs. However, taking a closer look at the items that Chris and Rob got right, we realize that they are different: Rob gave correct answers to the two easiest items, while Chris gave correct answers to two relatively complex items. This would of course require us to know the level of *difficulty* of each item, and would mean that while Chris gave incorrect answers to the easy items, he might still show some ability to read graphs. Thus, we would want the different scores to have 'meanings' to help us determine whether Chris was simply lucky (he *guessed* the answers), or whether he is in fact *able* to get the simpler items right, even though he didn't this time.

Imagine now that Rob, Jenny, and Chris take a second visualization literacy test. Rob gets a score of 3, Chris gets 4, and Jenny gets 10. We would infer that this test is easier, since the scores are higher. However, looking at the score intervals, we see that Jenny is 7 points ahead of Rob, whereas she was only 5 points ahead in the first test. If we were to truly measure abilities, we would want these intervals to be invariant. In addition, seeing that Chris' score is once again similar to Rob's (knowing that they both got the same items right this time) would lead us to think that they do in fact have similar abilities. We could then conclude that this test provides more *information* on lower abilities than the fist one, since it is able to separate Rob and Chris' scores.

Finally, imagine that all three examinees take a third test, and all get a score of 10. While we might be tempted to conclude that this test is visualization literacy-agnostic, it may simply be that its items are too easy, and not sufficiently *discriminant*.

One way of fulfilling all of these requirements is by using Item Response Theory (IRT) [277]. This is a model-based approach that does not use response data directly, but transforms them into estimates of a latent trait (*e.g.*, ability), which then serves as the basis of assessment. IRT models have been applied to tests in a variety of fields such as health studies, education, psychology, marketing, economics, social sciences (see [95]), and even graphicacy [268].

The core idea of IRT is that the performance of an examinee depends on both the examinee's ability and the item's difficulty; the goal is then to separate out these two factors. An important aspect of the approach is to project them onto the same scale—that of the latent trait. *Ability,* or standing on the latent trait, is derived from a pattern of responses to a series of test items; *item difficulty* is then defined by the 0.5 probability of success of an examinee with the appropriate ability. For example, an examinee with an ability value of 0 (0 corresponding to an average achiever) will have a 50% probability of giving a correct answer to an item with a difficulty value of 0, corresponding to an average level of difficulty.

IRT offers models for data that are dichotomous (*e.g.,* true/false responses) and polytomous (*e.g.,* responses on likert-like scales). Here, we focus on models for dichotomous data. These define the probability of success on an item i by the function:

$$p_i(\Theta) = c_i + \frac{1 - c_i}{1 + e^{-a_i(\Theta - b_i)}}$$

where Θ is an examinee's standing on a latent trait (*i.e.,* his or her ability), and a_i, b_i, and c_i are the characteristics of the item. The central characteristic is b, the *difficulty* characteristic; if $\Theta = b$, the examinee has a 0.5 probability of giving a correct answer to the item. Meanwhile, a is the *discrimination* characteristic. An item with a very high discrimination value basically sets a sharp threshold at $\Theta = b$: examinees with $\Theta < b$ have a probability of success of 0, and examinees with $\Theta > b$ have a probability of success of 1. Conversely, an item with a low discrimination value cannot clearly separate examinees. Finally, c is the *guessing* characteristic. It sets a lower bound for the extent to which an examinee will guess an answer. We have found c to be unhelpful, so we have set it to zero (no guessing) for our development.

Note that the value of each characteristic is not absolute for a given item: it is relative to the latent trait that the test is attempting to uncover. Therefore, it cannot be expected that the characteristics of identical items be exactly the same in different tests. For example, consider a simple numeracy test with two items: $10 + 20$ (item 1) and $17 + 86$ (item 2). It should be assumed that item 1 is easier than item 2. In other words, the difficulty characteristic of item 2 should be higher than that of item 1. Now if we add another item to the test, say 51×93 (item 3), the most difficult item in the previous version of the test (item 2) will no longer seem so difficult. However, it should still be more difficult than item 1. Thus, while individual characteristics may vary, the general order of difficulty should be preserved. The same goes for ability values (or *ability scores*). If they are to be compared between different tests, the measured latent trait must be the same.

Various IRT models for dichotomous data have been proposed. One is the *one-parameter logistic model* (1PL), which sets a to a specific value for all items, sets c to zero, and only considers the variation of b. Another is the *two-parameter logistic model* (2PL), which considers the variations of a and b, and sets c to zero. A third is the *three-parameter logistic model* (3PL), which considers variations of a, b, and c [54]. As such, 1PL and 2PL can be regarded as special cases of 3PL, where different item characteristics are assigned specific values. A last variant is the *Rasch model* (RM), which is a special case of 1PL, where $a = 1$[1].

Thus, IRT offers a way to evaluate the relevance of test items during a design phase (*e.g.,* how difficult items are, or how discriminant they are), and a way to measure examinees' abilities during an assessment phase. These two phases constitute the backbone of our method, which is why we stress that our approach will be successful only if an IRT model fits a set of empirically collected data. Furthermore, its accuracy will depend on how closely an IRT model describes the interaction between examinees' abilities and their responses, *i.e.,* how well the model describes the latent trait. Thus, different variants of IRT models should be tested initially to find the best fit. Finally, it should be mentioned that IRT models cannot be relied upon to 'fix' problematic issues in a test. Proper test design is still required.

1 For a complete set of references on the Rasch model, refer to http://rasch.org.

3.2 Foundations

In the approach we develop here, test items generally involve a three-part structure: 1) a *stimulus*, 2) a *task*, and 3) a *question*. The stimuli are the particular graphical representations used. Tasks are defined in terms of the visual operations and mental projections that an examinee should perform to answer a given question. While tasks and questions are usually linked, we emphasize this distinction because early piloting revealed that different 'orientations' of a question (*e.g.*, emphasis on particular visual aspects, or data aspects) could affect performance.

To identify possible factors that may influence the difficulty of a test item, we reviewed all the literacy tests that we could find which use graphs and charts as stimuli [[6], [22], [25], [58], [89], [94], [216], [268]]. Note that our goal is not to investigate the effect of these factors on item difficulty; we present them here merely as elements to be considered in the design phase. We identified four potential stimulus parameters: *number of samples*, *intrinsic complexity* (or variability) *of the data*, *layout*, and *level of distraction*. We also found six recurring task types: *extrema* (maximum and minimum), *trend*, *intersection*, *average*, and *comparison*. Finally, we distinguished three different question types: *perception* questions, *high-congruency* questions, and *low-congruency* questions. Each of these are described in the following subsections.

3.2.1 Stimulus parameters

In our survey, we first focused on identifying parameters to describe the graphical properties of a stimulus. We found four:

Number of samples—This refers to the number of graphically encoded elements in the stimulus. Among other things, the value of this parameter can impact tasks that require visual chunking [196].

Complexity—This refers to the *local* and *global variability* of the data. For example, a dataset of the yearly life expectancy in different countries over a 50 year time period has a low local variation (no dramatic 'bounces' between two consecutive years), and low global variation (a relatively stable, linear, increasing trend). In contrast, a dataset of the daily temperature in different countries over a year shows high local variation (temperatures can vary dramatically from on day to the other) and medium global variation (temperature generally rises an decreases only once during the year).

Layout—This refers to the structure of the graphical framework and its scales. Layouts can be *single* (*e.g.,* a two-dimensional Euclidian space), *superimposed* (*e.g.,* three axes for a 2dimensional encoding), or *multiple* (*e.g.,* several frameworks for a same visualization). Multiple layouts include *cutout charts* and *broken charts* [181]. Scales can be *single* (linear or logarithmic), *bifocal,* or *lense-like.*

Distraction—This refers to the graphical elements present in the stimulus that are not necessary for the task at hand. These are considered to be *distractors.* Correll *et al.* [140] have shown that even small variations in attributes of distractors can impact perception. However, here we simply use distraction in a Boolean way, *i.e.,* present or not.

3.2.2 Tasks

Next, we focused on identifying tasks that require only visual intelligence, *i.e.,* purely visual operations or mental projections on a graphical representation. We found six: *Maximum* (T1), *Minimum* (T2), *Variation* (T3), *Intersection* (T4), *Average* (T5), and *Comparison* (T6). All are standard benchmark tasks in Infovis. T1 and T2 consist in finding the maximum and minimum data points in the graph, respectively. T3 consists in detecting a trend, similarities, or discrepancies in the data. T4 consists in finding the point at which the graph intersected with a given value. T5 consists in estimating an average value. Finally, T6 consists in comparing different values or trends.

3.2.3 Congruency

Finally, we focused on identifying different types of questions. We found three: *perception* questions, and *high*- and *low-congruency* questions. Perception questions refer only to visual aspects of the display (*e.g.,* "what color are the dots?"). Conversely, congruent questions refer to semantic aspects of the data. The level of congruence is then defined by the *'replaceability' of the data-related terms in the question by perceptual terms.* A high-congruency question translates into a perceptual query simply by replacing data terms by perceptual terms (*e.g.*, "what is the highest value"/"what is the highest bar?"). A low-congruency question, in contrast, has no such correspondence (*e.g.,* "is A connected to B—in a matrix diagram"/"is the intersection between column A and row B highlighted?").

3.3 Application To Line Graphs

To illustrate our method, we first created two line graph tests—*Line Graphs 1* (LG1) and *Line Graphs 2* (LG2)—of slightly different designs, based on the principles described above. We then calibrated them using *Amazon's Mechanical Turk* (AMT).

3.3.1 Design Phase

3.3.1.1 Line Graphs 1: General Design

For our first test (LG1), we created a set of twelve items using different stimulus parameters and tasks. We hand-tailored each item based on an expected range of difficulty. Piloting had revealed that high variation in item dimensions led to incoherent tests (*i.e.*, IRT models did not fit the response data), implying that when factors vary too much within a test, additional abilities beyond those involved in basic visualization literacy are likely at play. Thus, we kept the number of varying factors low: only distraction and tasks varied. The test used four samples for the stimuli, and a single layout with single scales. A summary is given in TABLE 3.1.

Each item was repeated five times(2). The test was blocked by item, and all items and their repetitions were randomized to prevent carryover effects. We added an extra condition using a standard table at the beginning of each block to give examinees the opportunity to consolidate their understanding of the new question, and to separate out the comprehension stage of the *question-response* process believed to occur in cognitive testing [78]. The test was thus composed of 72 trials. In the following paragraphs, we describe other important design parameters we used in this test.

2 Early piloting had revealed that examinees would stabilize their search time and confidence after a few repetitions. In addition, repeated trials usually provide more robust measures as medians can be extracted (or means in the case of Boolean values).

LG1

Item ID	Task	Distraction
LG1.1	max	0
LG1.2	min	0
LG1.3	variation	0
LG1.4	intersection	0
LG1.5	average	0
LG1.6	comparison	0
LG1.7	max	1
LG1.8	min	1
LG1.9	variation	1
LG1.10	intersection	1
LG1.11	average	1
LG1.12	comparison	1

LG2

Item ID	Task	Congruency
LG1.1	max	high
LG1.2	min	high
LG1.3	variation	high
LG1.4	intersection	high
LG1.5	average	high
LG1.6	comparison	high
LG1.7	max	low
LG1.8	min	low
LG1.9	variation	low
LG1.10	intersection	low
LG1.11	average	low
LG1.12	comparison	low

TABLE 3.1: Designs of the Line Graphs 1 (LG1) and Line Graphs 2 (LG2) tests. Only varying dimensions are shown. Each item is repeated 6 times, beginning with a table condition (repetitions are not shown). Pink cells in the Item ID column indicate duplicate items in LG1 and LG2. Tasks with the same color coding are the same. Gray cells in the Distraction and Congruency columns indicate difference with white cells. The Distraction column uses a binary encoding: 0 = no distractors and 1 = presence of distractors.

Scenario—The *PISA 2012 Mathematics Framework* [216] emphasizes the importance of an understandable context for problem solving. The current test focuses on one's community, with problems set in a community perspective. To avoid the potential bias of *a priori* domain knowledge, the test was set within the following science-fiction scenario:

The year is 2813. The Earth is a desolate place. Most of mankind has migrated throughout the universe. The last handful of humans remaining on earth are now actively seeking another planet to settle on. Please help these people determine what the most hospitable planet is by answering the following series of questions as quickly and accurately as possible.

Data—The dataset we used had a low-local and medium-global level of variability. It presented the monthly evolution of unemployment in different countries between the years 2000 and 2008. Country names were changed to fictitious planet names listed in Wikipedia, and years were modified to fit the scenario.

Priming and Pacing—Before each new block of repetitions, examinees were primed with the upcoming graph type, so that the concepts and operations necessary for information extraction could be set up [228]. To separate out the time required to read questions, a specific pacing was given to each block. First, the question was displayed, along with a button labeled "Proceed to graph framework"; this led participants to the graphical framework with the appropriate title and labels. At the bottom of this was another button labeled "Display data," which displayed the full stimulus. As mentioned, to give examinees the opportunity to fully comprehend each question, every block began with a 'question comprehension' condition in which the data were shown in table form. This was intended to remove potential effects caused by the setup of high-level operations for solving a particular kind of problem. Finally, to make sure ability (and not capacity) was being tested, an 11 second (s) timeout was set for each repetition. This was based on the mean time required to answer the items in our pilot studies.

Response format—To respond, examinees were required to click on one of several possible answers, displayed in the form of buttons below the stimulus. In some cases, correct answers were not directly displayed. For example, certain values were not explicitly shown with labeled ticks on the graph's axes. This was done to test examinees' ability to make confident estimations (*i.e.*, to handle uncertainty [216]). In addition, although the stimuli used color coding to show different planets, the response buttons did not. This forced examinees to translate the answer found in the visual domain back into the data domain.

3.3.1.2 *Setup*

To calibrate our test, we administered it on AMT. While the validity of using this platform may be debated, due to lack of control over particular experimental conditions [175], we considered it best to perform our calibration using the results of a wide variety of people.

Participants—To our knowledge, no particular number of samples is recommended for IRT modeling. We recruited 40 participants who were required to have a 98% acceptance rate and a total of 1000 or more *Human Intelligence Tasks* (HITs) approved.

Coding—Six Turkers spent less than 1.5s on average reading and answering questions; they were considered as random clickers, and their results were removed from further analysis. All retained Turkers were native English speakers.

The remaining data were sorted according to item and repetition ID (assigned before randomization). Responses for the table conditions were removed. A score dataset (LG1s) was then created in accord with the requirements of IRT modeling: correct answers were scored 1 and incorrect answers 0. Scores for each set of item repetitions were then *compressed* by computing the rounded mean values. This resulted in a set of twelve dichotomous item scores for each examinee.

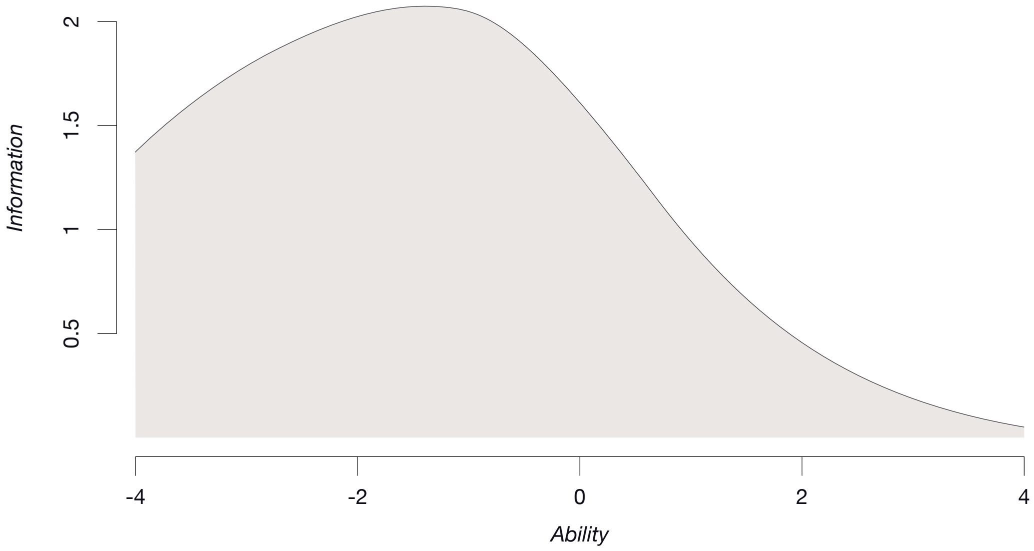
3.3.1.3 *Results*

The purpose of this calibration is to remove items that are unhelpful for distinguishing between low and high levels of visualization literacy. To do so, we need to: 1) check that the simplest variant of IRT models (*i.e.*, the Rasch model) fits the data, 2) find the best variant of the model to get the most accurate characteristic values for each item, and 3) assess the usefulness of each item.

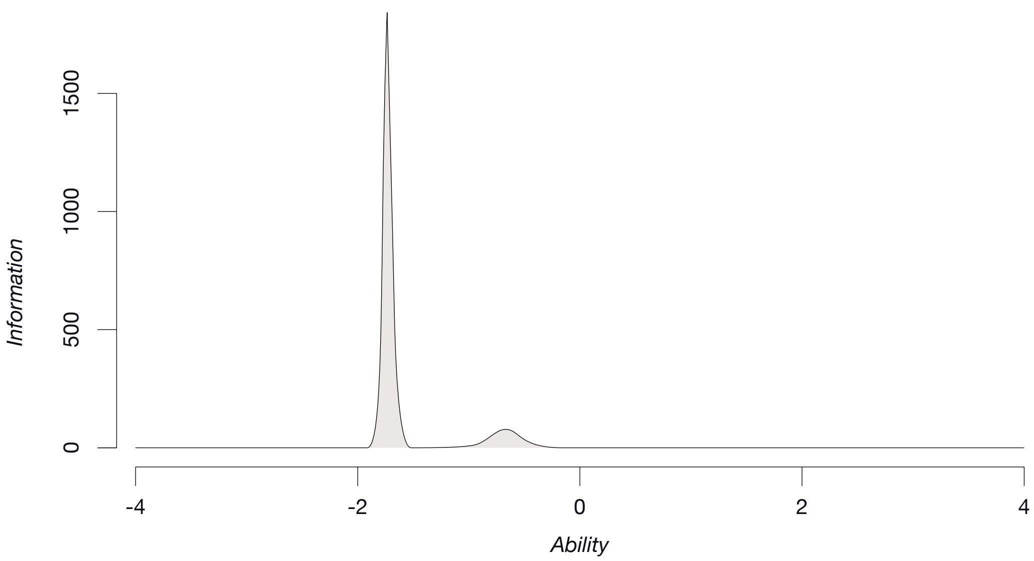
Checking the Rasch model—The Rasch model (RM) was first fitted to the score dataset. A 200 sample parametric Bootstrap goodness-of-fit test using Pearson's x^2 statistic revealed a non-significant p-value for LG1s ($p > 0.54$), suggesting an acceptable fit[(3)]. The *Test Information Curve* (TIC) is shown in FIGURE 3.2 (A). It reveals a near-normal distribution of test information across different ability levels, with a slight bump around −2, and a peak around −1. This means that the test provides more information about examinees with relatively low abilities (0 being the ability level of an average achiever) than about examinees with high abilities.

Finding the right model variant—Different IRT models, implemented in the *ltm* R package [232], were then fitted to LG1s. A series of pairwise likelihood ratio tests showed that the two-parameter logistic model (2PL) was most suitable. The new TIC is shown in FIGURE 3.2 (B).

3 For more information about this statistic, refer to [232].



(A) TIC of LG1s under RM.



(B) TIC of LG1s under 2PL.

FIGURE 3.2: Test Information Curves (TICs) of the score dataset of the first line graph test (LG1s) under the original constrained Rasch model (RM) (A) and the two-parameter logistic model (2PL) (B). The ability scale shows the Θ-values. The slight bump in the distribution of (A) can be explained by the presence of several highly discriminating items, as shown by the big spike in (B).

Assessing the usefulness of test items—The big spike in the TIC (FIGURE 3.2 (B)) suggests that several items with difficulty characteristics just above −2 have high discrimination values. This is confirmed by the very steep *Item Characteristic Curves* (ICCs—FIGURE 3.4 (A)) for items LG1.1, LG1.4, and LG1.9 ($a > 51$), and can explain the slight distortion in FIGURE 3.2 (A).

The probability estimates revealed that examinees with average abilities have a 100% probability of giving a correct answer to the easiest items (LG1.1, LG1.4, and LG1.9), and a 41% probability of giving a correct answer to the hardest item (LG1.11). However, the fact that LG1.11 has a relatively low discrimination value ($a < 0.7$) suggests that it is not very effective for separating ability levels.

3.3.1.4 *Discussion*

IRT modeling appears to be a solid approach for calibrating our test design. Our results (FIGURE 3.2) show that LG1 is useful for differentiating between examinees with relatively low abilities, but not so much for ones with high abilities. The slight bump in the distribution of the TIC (FIGURE 3.2 (A)) suggests that several test items are quite effective for separating ability levels around −2. This is confirmed by the spike in FIGURE 3.2 (B), which indicates the presence of highly discriminating items. Overall, both Test Information Curves reveal that the test is best suited for examinees with relatively low abilities, since most of the information it provides concerns ability levels below zero. In addition, FIGURE 3.4 (A) reveals that several items in the test have identical difficulty and discrimination characteristics. Some of these could be considered for removal, as they provide only redundant information. Similarly, item LG1.11, which has a low discrimination characteristic, could be dropped, as it is less effective than others.

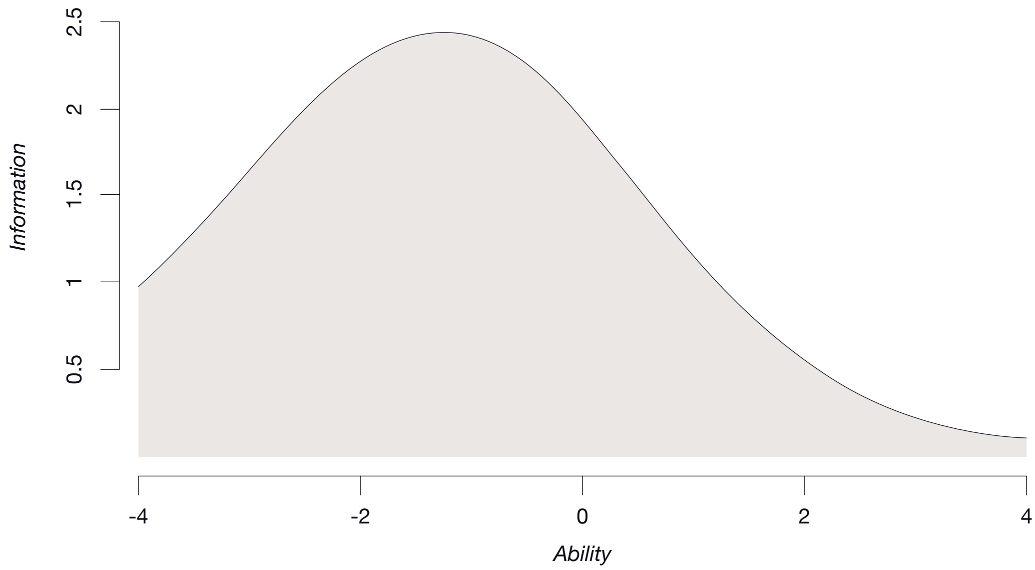
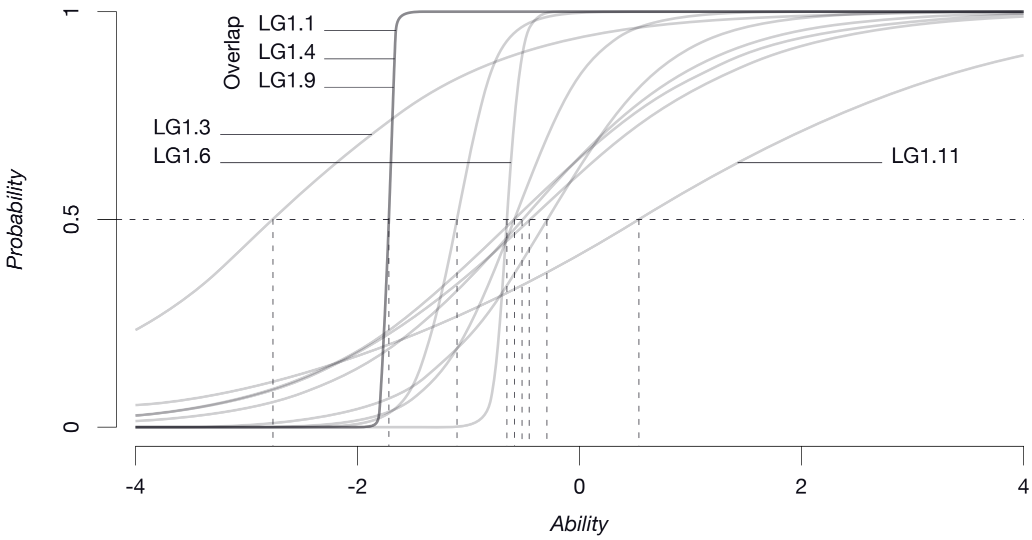


FIGURE 3.3: Test Information Curve of the score dataset of the second line graph test (LG2s) under the original constrained Rasch model. The test information is normally distributed.

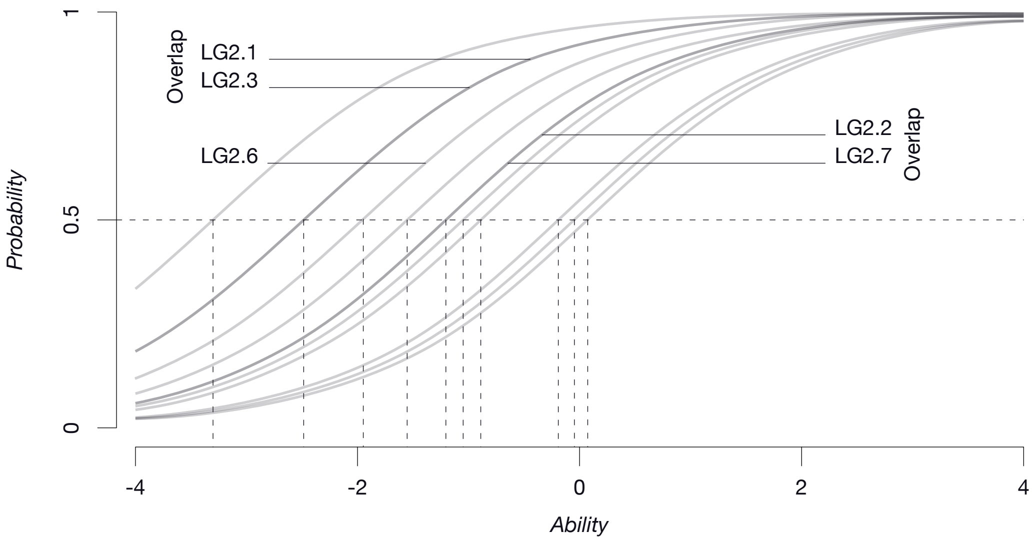
3.3.1.5 *Line Graphs 2: General Design*

For our second line graph test (LG2), we also created twelve items, with varying factors restricted to question congruency and tasks (see TABLE 3.1). The test used four samples for the stimuli, and a single layout with single scales. The same scenario, dataset, pacing, and response format as for LG1 were kept, as well as the five repetitions, the question comprehension condition, and the 11s timeout. As such, six items in this test were identical to items in LG1 (see pink cells in TABLE 3.1). This was done to ensure that the order of item difficulty would remain consistent across the different tests.

The calibration was again conducted on AMT. 40 participants were recruited; the work of three Turkers was rejected, for the same reason as before.



(A) ICCs of LG1s under 2PL.



(B) ICCs of LG2s under RM.

FIGURE 3.4: Item Characteristic Curves (ICCs) of the score datasets of the first line graph test (LG1s) under the original the two-parameter logistic model (A), and of the second line graph test (LG2s) under the constrained Rasch model (B). The different curve steepnesses in (a) are due to the fact that 2PL computes individual discrimination values for each item, while RM sets all discrimination values to 1.

3.3.1.6 *Results and Discussion*

Our analysis was driven by the requirements listed above. Data were sorted and encoded in the same way as before, and a score dataset for LG2 was obtained (LG2s). RM was fitted to the score dataset, and the goodness-of-fit test revealed an acceptable fit ($p > 0.3$). The pairwise likelihood ratio test showed that RM was the best of all possible IRT models. The Test Information Curve (FIGURE 3.3) is normally distributed, with a peak around -1. This indicates that like our first line graph test, LG2 is best suited for examinees with relatively low abilities.

The Item Characteristic Curves of both tests were then compared. While it cannot be expected that identical items have the exact same characteristics, their difficulty order should remain consistent (see Section 3.1.3). FIGURE 3.4 shows some slight discrepancies for items 1, 3, and 6 between the two tests. However, the fact that item LG1.3 is further to the left in FIGURE 3.4 (A) is misleading. It is due to the extremely high a-values of items LG1.1 and LG1.4. Thus, while their b-values are slightly higher than that of LG1.3, the probability of success of an average achiever is higher for these items than it is for LG1.3 ($1 > 0.94$). Furthermore, the difficulty characteristics of LG1.3 and LG2.3 are very similar ($0.94 \approx 0.92$). Therefore, the only exception in the ordering of item difficulties is item 6, which is estimated to be more difficult than item 2 in LG1, and not in LG2.

This suggests that LG1 and LG2 cover the same latent trait, *i.e.,* ability to read line graphs. To examine this, we equated the test scores using a *common item equation* approach. RM was fitted to the resulting dataset, the goodness-of-fit test showed an acceptable fit, and 2PL provided best fit. Individual item characteristics were generally preserved, with the exception of item 6, which, interestingly, ended up with characteristics very similar to those of item 2. This confirms that the two tests cover the same latent trait. Thus, although individual characteristics are slightly altered by the equation (*e.g.,* item 6), items in LG1 can safely be transposed to LG2, without hindering the overall coherence of the test, and *vice-versa.*

3.3.2 Assessment Phase

Having shown that our test items have a sound basis in theory, we now turn to the assessment of visualization literacy. While a standard method would simply sum up the correct responses, our method considers each response individually, with regard to the difficulty of the item it was given for. To make this assessment, we inspected the ability scores derived from the fitted IRT models. These scores represent examinees' standings (Θ) on the latent trait, and correspond to a unique response pattern. They have great predictive power as they can determine an examinee's probability of success on items that s/he has not completed, provided that these items follow the same latent variable scale as other test items. As such, ability scores are perfect indicators for assessing visualization literacy.

LG1 revealed **27 different ability scores**, ranging from −1.85 to 1. The distribution of these scores was near-normal, with a slight bump around −1.75. **40.7%** of participants were above average (*i.e.,* $\Theta > 0$), and the mean was **−0.27**. LG2 revealed **33 different ability scores**, ranging from −1.83 to 1.19. The distribution was also near-normal, with a bump around −1. **39.4%** of participants were above average, and the mean was **−0.17**.

These results show that the means are close to zero, and the distributions near-normal. This suggests that most Turkers, while somewhat below average in visualization literacy for line graphs, have fairly standard abilities. While it should be interesting to develop broader ranges of item complexities for the line graph stimulus (by using the common item equation approach), thus extending the psychometric quality of the tests, we consider LG1 and LG2 to be sufficient for our current line of research. Furthermore, we believe that these low levels of difficulty reflect the general simplicity of, and massive exposure to, line graphs.

BC

Item ID	Task	Samples
LG1.1	max	10
LG1.2	min	10
LG1.3	variation	10
LG1.4	intersection	10
LG1.5	average	10
LG1.6	comparison	10
LG1.7	max	20
LG1.8	min	20
LG1.9	variation	20
LG1.10	intersection	20
LG1.11	average	20
LG1.12	comparison	20

SP

Item ID	Task	Distraction
LG1.1	max	0
LG1.2	min	0
LG1.3	variation	0
LG1.4	intersection	0
LG1.5	average	0
LG1.6	comparison	0
LG1.7	max	1
LG1.8	min	1
LG1.9	variation	1
LG1.10	intersection	1
LG1.11	average	1
LG1.12	comparison	1

TABLE 3.2: Design of the Bar Charts (BC) and Scatterplots (SP) tests. Only varying dimensions are shown. Each item is repeated 6 times, beginning with a table condition (repetitions are not shown). Tasks with the same color coding are the same. Gray cells in the Samples and Distraction columns indicate difference with white cells. The Distraction column uses a binary encoding: 0 = no distractors and 1 = presence of distractors.

3.4 Extensions

To see whether our method also applies to other visualizations, we created two additional tests: one for *Bar Charts* (BC) and one for *Scatterplots* (SP).

3.4.1 Design Phase

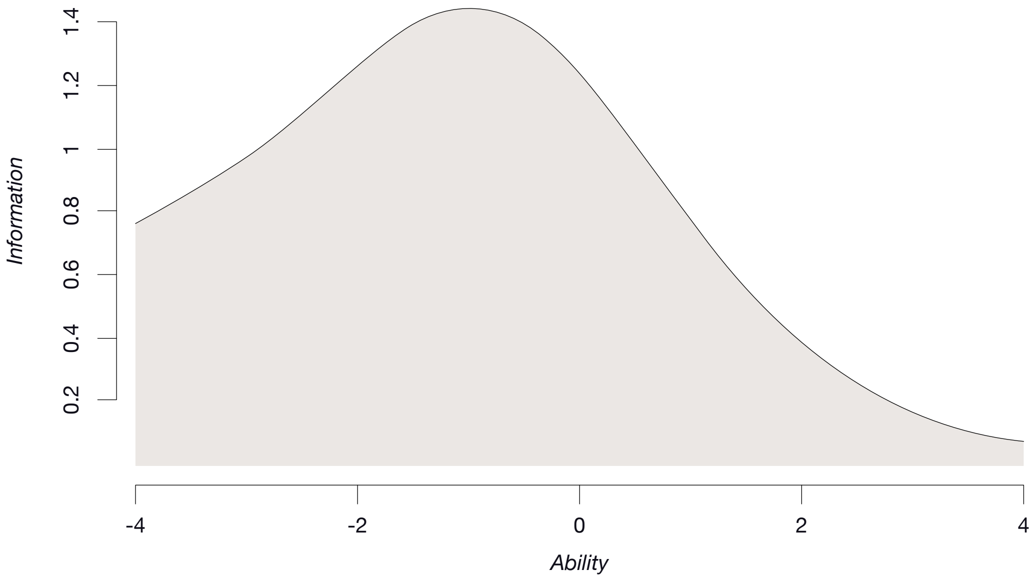
3.4.1.1 *Bar Charts: General Design*

Like LG1 and LG2, the design of our bar charts test (BC) was based on the principles described in Section 3.2. We created twelve items, with varying factors restricted to number of samples and tasks (see TABLE 3.2). The same scenario, pacing, response format, repetitions, question comprehension condition, and 11s timeout were kept. The dataset presented life expectancy in various countries, with country names again changed to fictitious planet names. The only difference with the factors used earlier (apart from the stimulus) involved the variation task, which is essentially a trend detection task. Bar charts are sub-optimal for determining trends (as they are meant for comparing data across discrete categories [97]), so this task was replaced by a *global similarity detection* task, as done in [94] (*e.g.,* "Do all the bars have the same value?").

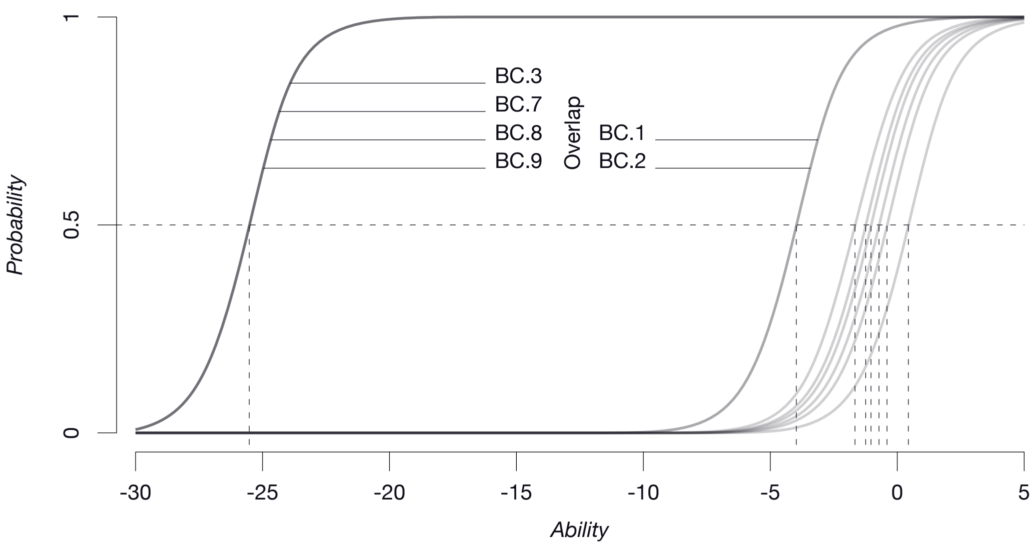
The calibration was again conducted on AMT. 40 participants were recruited; the work of six Turkers was rejected, for the same reason as before.

3.4.1.2 *Results and Discussion*

Our analysis was driven by the same requirements as for the line graph tests. Data were sorted and encoded in the same way, resulting in a score dataset for BC (BCs). RM was first fitted to BCs; the goodness-of-fit test revealed an acceptable fit ($p > 0.37$), and the likelihood test proved that it fit best. However, the Test Information Curve (FIGURE 3.5 (A)) is not normally distributed. This is due to the presence of several extremely low dif-

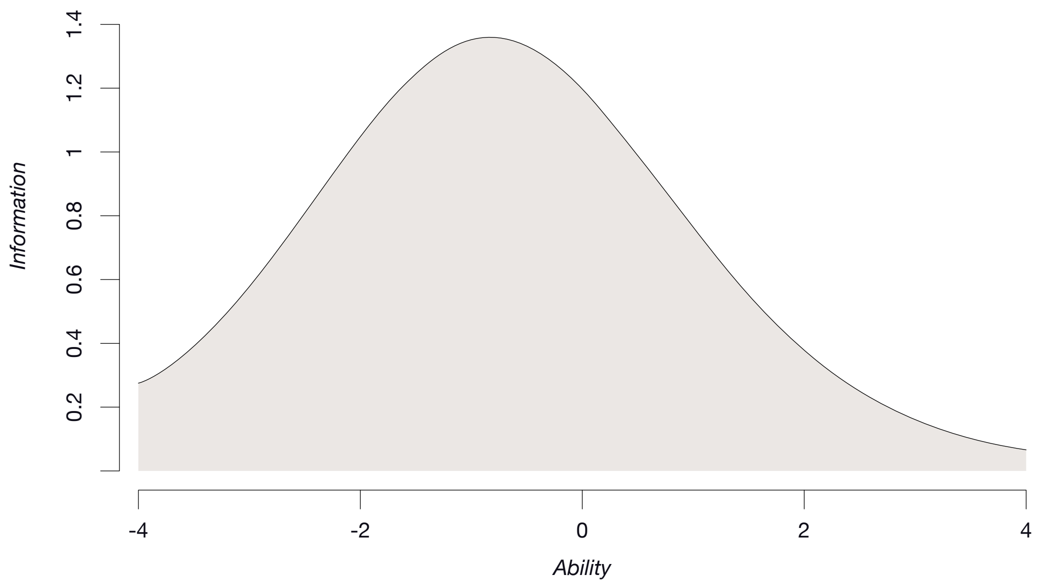


(A) TIC of BCs under RM.

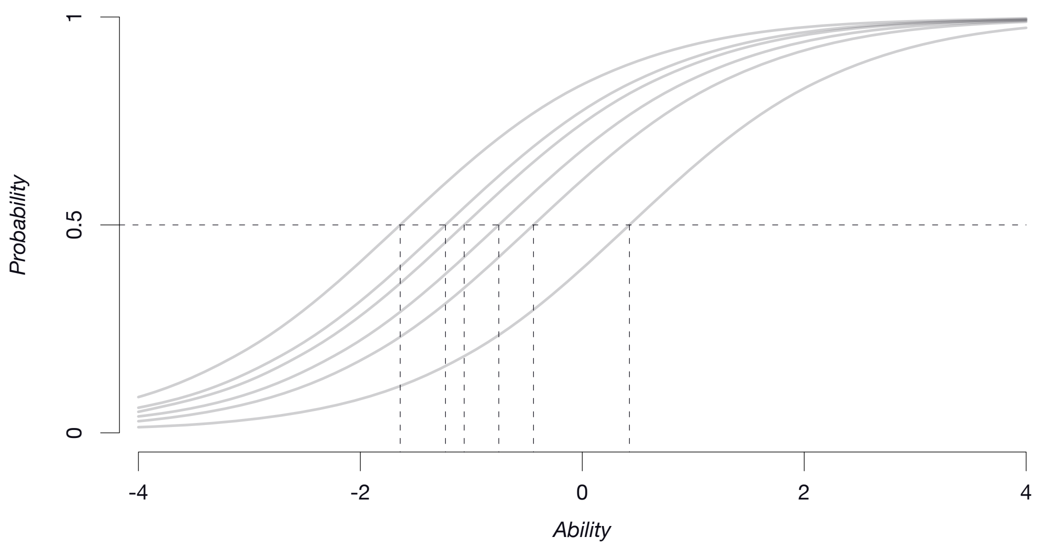


(B) ICCs of BCs under RM.

FIGURE 3.5: Test Information Curve (A) and Item Characteristic Curves (B) of the score dataset of the bar chart test under the constrained Rasch model. The TIC in (A) is not normally distributed because of several very low difficulty items, as shown by the curves to the far left of (B).



(A) TIC of the subset of BCs under RM.



(B) ICCs of the subset of BCs under RM.

FIGURE 3.6: Test Information Curve (A) and Item Characteristic Curves (B) of the subset of the score dataset of the bar chart test under the constrained Rasch model. The subset was obtained by removing the very low difficulty items shown in FIGURE 3.5 (B).

ficulty (*i.e.*, easy) items (BC.3, BC.7, BC.8, and BC.9; $b = -25.6$), as shown in FIGURE 3.5 (B). Inspecting the raw scores for these items revealed a 100% success rate. Thus, they were considered too easy, and were removed. Similarly, items BC.1 and BC.2 (for both, $b < -4$) were also removed.

To check the coherence of the resulting subset of items, RM was fitted again to the remaining set of scores. Goodness-of-fit was maintained ($p > 0.33$), and RM still fitted best. The new TIC (FIGURE 3.6 (A)) is normally distributed, with a peak around -1. This indicates that like our line graph tests, this subset of BC is best suited for examinees with relatively low abilities.

3.4.1.3 *Scatterplots: General Design*

For our scatterplot test (SP), we once again created twelve items, with varying factors restricted to distraction and tasks (see TABLE 3.2). The same scenario, pacing, response format, repetitions, and question comprehension condition were kept. The dataset presented levels of adult literacy by expenditure per student in primary school in different countries, with country names again changed to fictitious planet names. Slight changes were required for some of the tasks, since scatterplots use two spatial dimensions (as opposed to bar charts and line graphs). For example, stimuli with distractors in LG1 only required examinees to focus on one of several samples; here, stimuli with distractors could either require examinees to focus on a single datapoint or on a single dimension. Finally, we had initially expected that SP would be more difficult, and items would require more time to complete. However, a pilot study showed that the average response time per item was again roughly 11s. Therefore, the 11s timeout condition was kept.

The calibration was again conducted on AMT. 40 participants were recruited; the work of one Turker was not kept because of technical (logging) issues.

3.4.1.4 *Results and Discussion*

Our analysis was once again driven by the same requirements as before. The same sorting and coding was applied to the data, resulting in the score dataset SPs. The fitting procedure was then applied, revealing a good fit for RM ($p = 0.6$), and a best fit for 2PL.

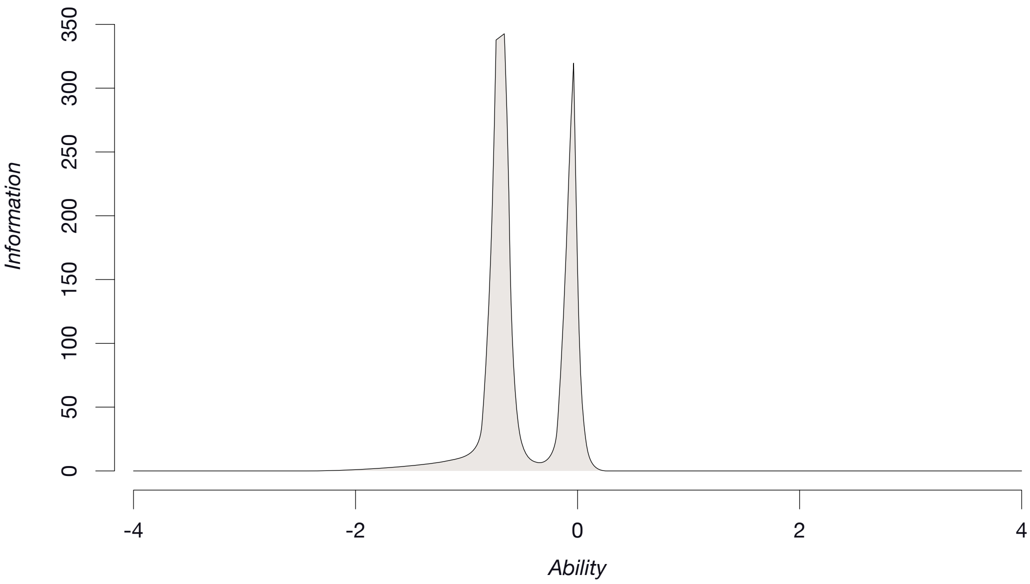
The Test Information Curve (FIGURE 3.7 (A)) shows the presence of several highly discriminating items around $b \approx -1$ and $b \approx 0$. The Item Characteristic Curves (FIGURE 3.7 (B)) confirm that there are three (SP.6, SP.8, and SP.10; $a > 31$). However, they also show that two items (SP.3, and SP.11) have quite low discrimination values ($a < 0.6$). Here, we set a threshold for $a > 0.8$. Thus, items SP.3 and SP.11 were removed. The resulting subset of 10 items' scores was fitted once again. RM fitted well ($p = 0.69$), and 2PL fitted best. The different curves of the subset are plotted in FIGURE 3.8. They show a good amount of information for abilities that are slightly below average (FIGURE 3.8 (A)), which indicates that the subset of SP is once again best suited for examinees with relatively low abilities.

3.4.2 Assessment phase

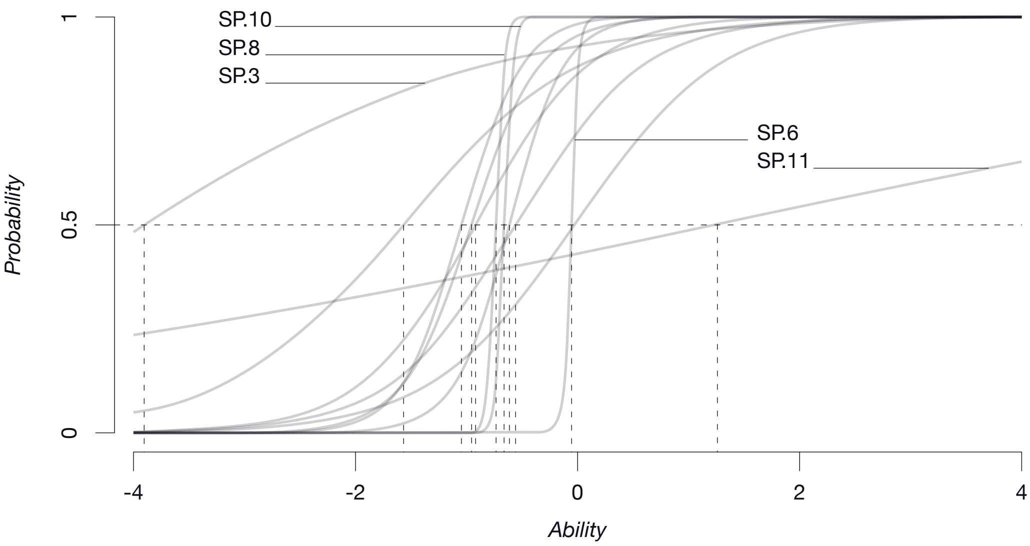
Here again, we inspected the Turkers' ability scores. Only the items retained at the end of the design phase were used.

BC revealed **21 different ability scores**, ranging from −1.75 to 0.99. The distribution of these scores was near-normal, with a slight bump around −1.5, and the mean was **−0.39**. However, only **14.3%** of participants were above average. SP revealed **23 different ability scores**, ranging from −1.72 to 0.72. However, the distribution here was not normal. **43.5%** of participants were above average, and the median was **−0.14**.

These results show that the majority of recruited Turkers had somewhat below average levels of visualization literacy for bar charts and scatterplots. The very low percentage of Turkers above average in BC led us to reconsider the removal of items BC.1 and BC.2, as they were not truly problematic. After reintegrating them in the test scores, 21 ability scores

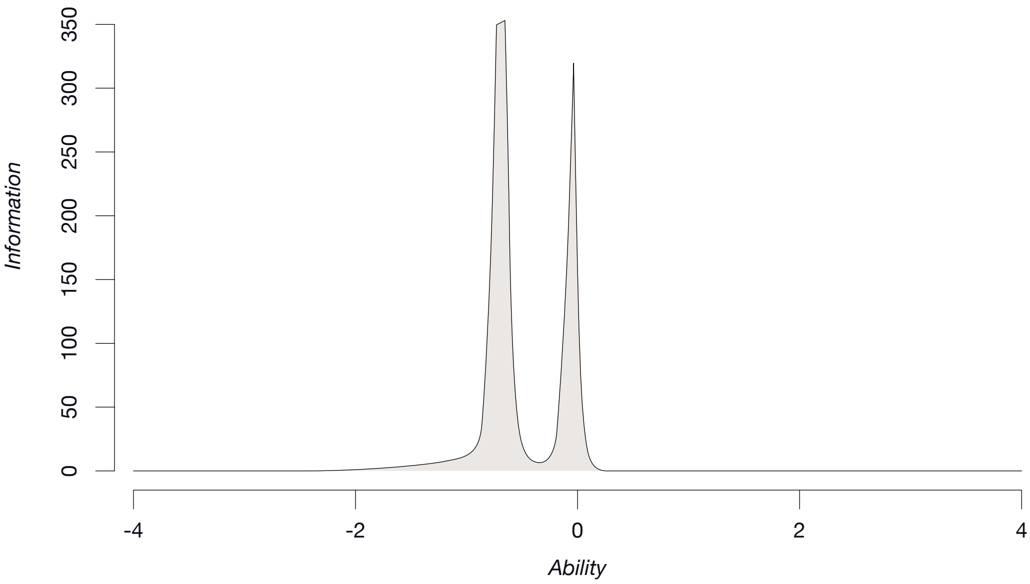


(A) TIC of SPs under 2PL.

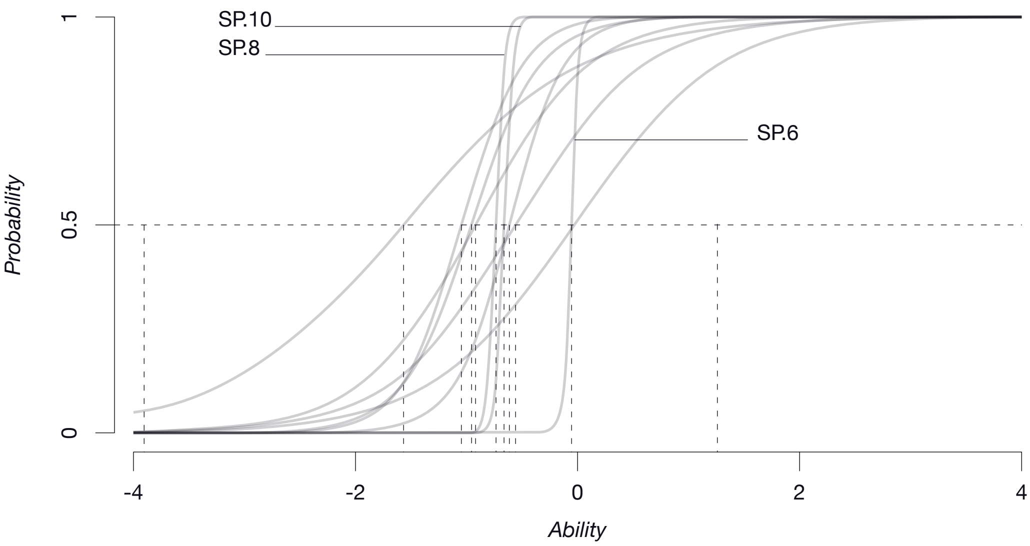


(B) ICCs of SPs under 2PL.

FIGURE 3.7: Test Information Curve (A) and Item Characteristic Curves (B) of the score dataset of the scatterplot test under the two-parameter logistic model. The TIC (A) shows that there are several highly discriminating items, which is confirmed by the very steep curves in (B). In addition, (B) shows that there are also a two poorly discriminating items, represented by the very gradual slopes of items SP.3 and SP.11.



(A) TIC of the subset of SPs under 2PL.



(B) ICCs of the subset of SPs under 2PL.

FIGURE 3.8: Test Information Curve (A) and Item Characteristic Curves (B) of the subset of the score dataset of the scatterplot test under the two-parameter logistic model. The subset was obtained by removing the poorly discriminating items shown in FIGURE 3.7 (B).

were observed, ranging from −1.67 to 0.99, and 42.8% of participants were above average. This seemed more convincing. However, this important difference illustrates the relativity of these values, and shows how important it is to properly calibrate the tests during the design phase.

Finally, we did not attempt to equate these tests, since—unlike LG1 and LG2—we ran them independently without any overlapping items. To have a fully comprehensive test, *i.e.*, a generic test for visualization literacy, intermediate tests are required where the stimulus itself is a varying factor. If such tests prove to be coherent (*i.e.*, if IRT models fit the results), then it should be possible to assert that visualization literacy is a general trait that allows one to understand any kind of graphical representation. Although we believe that this ability varies with exposure and habit of use, a study to confirm it is outside of the scope of this current work.

3.5 Fast, Effective Testing

If these tests are to be used as practical ways of assessing visualization literacy, the process must be sped up, both in the administration of the tests and in the analysis of the results. While IRT provides useful information on the quality of tests and on the ability of those who take them, it is quite costly, both in time and in computation. This must be changed.

In this section, we present a way in which the tests we have developed in the previous sections can be optimized to be faster, while still maintaining their effectiveness.

3.5.1 Test Administration Time

As we have seen, several items can be removed from the tests, while keeping good psychometric quality. However, this should be done carefully, as some of these items may provide useful information (like in the case of BC.1 and BC.2, Section 3.4.2).

We first removed LG1.11 from LG1, as its discrimination value was < 0.8 (see Section 3.4.1.4). We then inspected items with identical difficulty and discrimination characteristics, represented by overlapping ICCs' (see FIGURE 3.4). These were prime candidates for removal, since they provide only redundant information. There was one group of overlapping items in LG1 ([LG1.1, LG1.4, LG1.9]), and two in LG2 ([LG2.1, LG2.3], [LG2.2, LG2.7]). For each group, we kept only one item. Thus LG1.1, LG1.4, LG2.3, and LG2.7 were dropped.

We reintegrated items BC.1 and BC.2 to BC, as they proved to have a big impact on ability scores (Section 3.4.2). The subset of SP created at the end of the design phase was kept, and no extra items were removed.

RM was fitted to the newly created subsets of LG1, LG2, and BC; the goodness-of-fit test showed acceptable fits for all ($p > 0.69$ for LG1, and $p > 0.3$ for both LG2 and BC). 2PL fitted best for LG1, and RM fitted best for both LG2 and BC.

We conducted a *post-hoc* analysis to see whether the number of item repetitions could be reduced (first to three, then to one). Results showed that RM fitted all score datasets using three repetitions. However, several examinees had lower scores. In addition, while BC and SP showed similar amounts of information for the same ability levels, the three very easy items in BC (*i.e.,* BC.3, BC.7, BC.8, and BC.9) were no longer problematic. This suggests that several participants did not get a score of 1 for these items, and confirms that, for some examinees, more repetitions are needed. Results for one-repetition-tests showed that RM no longer fitted the scores of BC, suggesting that unique repetitions are noisy. Therefore, we decided to keep the five repetitions.

In the end, the redesign of LG1 contained 9 items (with a ≈ 10 min completion time—see Appendix E), the redesigns of LG2 and SP contained 10 items (11 min—see Appendix F and Appendix H, respectively), and the redesign of BC contained 8 items (≈ 9 min—see Appendix G).

3.5.2 Analysis Time and Computational Power

To speed up the analysis, we first considered setting up the procedure we had used in R on a server. However, this solution would have required a lot of computational power, so we dropped it.

Instead, we chose to tabulate all possible ability scores for each test. An interesting feature of IRT modeling is that it can derive ability scores from unobserved response patterns (*i.e.,* patterns that do not exist in the empirical data), as well as from partial response patterns (*i.e.,* patterns with missing values). Consequently, we generated all the $2^{n_i} - 1$ possible patterns for each test, where n_i is the number of items in a test. This resulted in 511 patterns for LG1, 1023 for both LG2 and SP, and 255 for BC. We then derived the different ability scores that could be obtained in each test.

To ensure that removing certain test items did not greatly affect the ability scores, we computed all the scores for the full LG1 and LG2 tests, and compared them to the ones previously obtained. We found some small differences in the upper and lower bounds of ability, but these were considered negligible, since our tests were not designed for fine distinction between very low abilities or high abilities. We also tested the impact of refitting the IRT models after item removal. For this, we repeated the procedure using partial response patterns for LG1 and LG2, *i.e.*, we replaced the dichotomous response values for the items considered for removal by *not available* (NA) values. The scores were exactly the same as the ones obtained with our already shortened and refitted tests, which proves they can be trusted.

Finally, we integrated all ability scores and their corresponding response patterns into the web-based, shortened versions of the tests, to make them readily available. This way, by administering our online tests, researchers can have direct access to participants' levels of visualization literacy. Informed decisions can then be made as whether to keep these people for further studies or not. All four tests are accessible at http://peopleviz. gforge.inria.fr/trunk/vLiteracy/home/.

3.6 Methodology Guidelines

As the preceding sections have shown, we have developed and validated a fast and effective method for assessing visualization literacy. This section summarizes the major steps, written in the form of easy takeaway guidelines.

3.6.1 Initial Design

1. **Pay careful attention to the design of all three components of a test item**, *i.e.*, stimulus, task, and question. Each can influence item difficulty, and too much variation may fail to produce a coherent test—as was seen in our pilot studies.
2. **Repeat each item several times**. We did 5 repetitions + 1 'question comprehension' condition for each item. This is important as repeated trials provide more robust measures. Ultimately, it may be feasible to reduce the number of repetitions to 3[4], although our results show that this can be problematic (Section 3.5.1).
3. **Use a different—and ideally, non-graphical—representation for question comprehension**. We chose a table condition. While our present study did not focus on proving its essentialness, we believe that this attribute is important.
4. **Randomize the order of items and of repetitions**. This is common practice in experiment design, having the benefit of preventing carryover effects.
5. Once the results are in, **sort the data according to item and repetition ID**, **remove the data for the**

4 The number of repetitions should be odd, so as to not end up with a mean score of 0.5 for an item.

question comprehension condition, and encode examinees' scores in a dichotomous way, *i.e.*, 1 for correct answers and 0 for incorrect answers.

6. **Calculate the mean score for all repetitions of an item and round the result**. This will give a finer estimate of the examinee's ability since it erases one-time errors which may be due to lack of attention or to clicking on the wrong answer by mistake.
7. **Begin model fitting with the Rasch model**. RM is the simplest variant of IRT models. If it does not fit the data, other variants will not either. Then **Check the fit of the model**. Here we used a 200 sample parametric Bootstrap goodness-of-fit test using Pearson's χ^2 statistic. To reveal an acceptable fit, the returned p-value should not be statistically significant ($p > 0.05$). In some cases (like in our pilot studies), the model may not fit. Options here are to inspect the χ^2 p-values for pairwise associations, or the two- and three-way χ^2 residuals, to find problematic items[5].
8. **Determine which IRT model variant best fits the data**. A series of pairwise likelihood ratio tests can be used for this. If several models fit, it is usually good to go with the model that fits best. Our experience showed that such models were most often RM and 2PL.
9. **Identify potentially useless items**. In our examples of LG1 and SP, certain items had low discrimination characteristics. These are not very effective for separating ability levels, and can be removed. In cases like the one for BC, items may also simply be too easy. Before removing them permanently, however, it is advised to check their impact on ability scores.

5 For more information, refer to [232].

Finally, it is important the model be refitted at this stage (reproducing steps 7 and 8), as removing these items may affect examinee and item characteristics.

3.6.2 Final Design

10. **Identify overlapping items and remove them**. If the goal is to design a short test, such items can safely be removed, as they provide only redundant information (see Section 3.5.1).
11. **Generate all $2^{n_i}-1$ possible score patterns**, where n_i is the number of retained items in the test. These patterns represent series of dichotomous response values for each test item.
12. **Derive the ability scores from the model**, using the patterns of responses generated in step 11. These scores represent the range of visualization literacy levels that the test can assess.
13. **Integrate the ability scores into the test** to make fast, effective, and scalable estimates of people's visualization literacy.

3.7 Conclusion

This chapter has focused on the **literacy cost**, and has addressed the following research question:

Q1: How can a designer know the level of understanding an audience has of different visual representations of data?

I have presented a method for assessing visualization literacy, which I developed with Ronald A. Rensink, Enrico Bertini, and Jean-Daniel Fekete. This method is based on a principled set of considerations, and in particular on Item Response Theory to allow a separation of the effects of item difficulty and examinee ability. Our motivation was to make a series of fast, effective, and reliable tests which researchers could use to detect participants with low visualization literacy abilities before conducting online studies. We have shown how these tests can be tailored to get immediate estimates of examinee's levels of visualization literacy.

While we have adopted an evaluation approach, we see several ways in which this work can help designers know the level of understanding an audience has of different visual representations of data (**Q1**). First, the tests developed here can be used during participatory design sessions to assess a targeted audience's level of visualization literacy. This can help adapt design requirements to best suit their skills. Second, we believe that our definition of *congruency* (see Section 3.2.3) can be extended to describe different types of visual representations, based on how easily they allow to translate questions set in the data-domain into visual queries. We hypothesize that initially providing casual audiences with highly-congruent visualizations can help them understand the concept of visual mapping, and can increase their visualization literacy. By building up a feeling of self-efficacy, this can possibly lead users to put the effort into understanding less congruent visualizations. A way to apply this would be to design a narrative visualization that uses an interactive slideshow format in which highly-congruent visualizations are introduced in the first slides, leading up to less-congruent visualizations, which users can choose to switch to when they feel confident and autonomous.

In line with this, another interesting extension to the work presented in this chapter would be to identify an *indirect* measure of visualization literacy, *i.e.,* a proxy, which would indicate how well a user understands a given visual representation, and could be automatically detected by online tracing systems. This could help automatically guide users through different visual representations in accord with their skills. While our current tests have been optimized to be as easy and quick to administer as possible, they are still inconvenient outside of planned studies, *i.e.,* deploying them over the web as an initial component of an information visualization of open data seems impractical. Taking a test can be stressful, and most information seekers are unlikely to be willing to do so just to get access to a website—they are most likely to simply bounce away.

Thus, we acknowledge that this work is but a small step into the realm of visualization literacy, which is why we have made our tests available on GitHub for versioning [101]. These can serve as a basis for establishing the suitability of our method for other kinds of representations (*e.g.,* parallel coordinates, node link diagrams, starplots, *etc.*), and possibly for other purposes. In contexts like a classroom evaluation, the tests could be longer, and broader assessments of visualization literacy could be made. This would imply further exploration of the design parameters proposed in Section 3.2. Evaluating the impact of these parameters on item difficulty should also prove interesting. Ultimately, we hope that this work will serve as a foundation for further research into visualization literacy.

Chapter 4

Designing for Context-Interpretation: Bridging the Gap Between Infovis and Visual Communication

In October 2012, Alberto Cairo posted a chapter from his book *The Functional Art* on Peachpit.com, in which he discusses the challenges of infographics design from a media perspective [31]. He mentions the goals of a *good* visualization are to present information, while allowing users to explore that information; and he describes an information graphic as both "a tool for the designer to communicate with readers, and a tool for readers to analyze what's being presented to them." He then concludes by addressing visualization designers, and stating: "it is crucial to remember our priorities as visual communicators." But what are the priorities of visual communication? And how can a graphic communicate with a reader? Although the language of visualization is rich, it is mostly very abstract. This makes it difficult to interpret at a glance what data are about; and as I have mentioned in the previous chapters, this **context-interpretation cost** can be another important drawback for casual audiences.

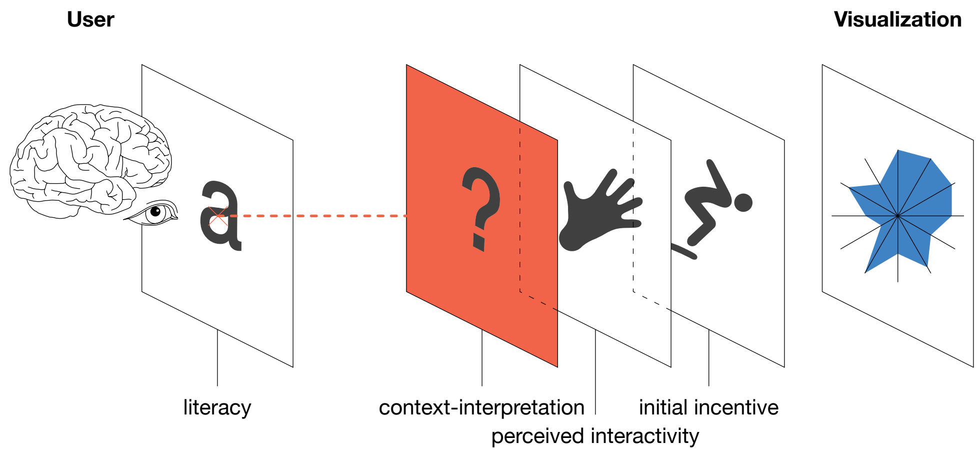


FIGURE 4.1: The **context-interpretation cost**.

This chapter is based on a published poster paper entitled *The CO2 Pollution Map: Lessons Learned from Designing a Visualization that Bridges the Gap between Visual Communication and Information Visualization* [123], so any use of the term "we" refers to myself and Jean-Daniel Fekete. It focuses on the **context-interpretation cost** by addressing the following research question:

Q2: How can visualizations be designed to help people interpret their context, *i.e.*, the semantic nature of the data they present?

To answer this, we present the design of a visualization that takes inspirations from the disciplines of graphic and motion design, and illustrates how transposing design considerations used in other fields can impact visualization design choices. Traditionally, infovis design has been driven by analytic needs, which emphasize visual clarity to avoid cognitive biases in the interpretation of data. Unfortunately, visual clarity does not always enforce message clarity.

As information visualization is now often used as a medium for journalism, designers need to find ways to optimize the communicativeness of visualizations. An interesting way to approach this is to investigate how graphic and motion designers use other visual representations to communicate information. Essentially, their goal is to harness a viewer's attention before attracting him/her into the 'complexity' of an image. They create intricate visual languages based on compelling forms that intend to engage the viewer emotionally with both the medium and its content.

Our approach is inspired by these practices, and focuses on the design of the *CO2 Pollution Map*—a visualization that uses a particle system (a computer graphics technique traditionally used in motion design to create *fuzzy* objects like smoke) to present viewers with a visual metaphor for pollution. This metaphor is intended to both encode data, and help viewers rapidly interpret what they are about. From this design, we derived a framework for thinking about visualization design from a visual communication perspective.

As such, the main contributions of this chapter are as follows:

* a description of the implementation of the *CO2 Pollution Map*;
* a set of considerations derived from our design; and
* a novel framework for thinking about visualization design.

Note that we adopt a design approach here, as our immediate motivation for this work is to find a way to rethink how information visualizations communicate the topic of a dataset to a viewer—without using additional text or embellishments. This requires bridging the gap between information visualization design and visual communication. However, we stress that the approach developed here should be considered carefully, as it may impede on the analytic aspects of infovis. As Cairo puts it, "graphics should not simplify messages. They should clarify them" [31].

This chapter is organized in the following way. It begins with a short background section that extends both Section 2.2.2, and Section 2.3.2.1; it presents three levels of visual communication, inspired by the work of graphic design practitioners, and emphasizes once again the importance of semantic design. In Section 4.2, we describe the implementation of a *CO2 Pollution Map*. Section 4.3 discusses the benefits and limitations of our design, and presents some early results of the deployment of the visualization. Finally, Section 4.4 introduces the framework we derived from this work, which intends to bridge the gap between visual communication and information visualization design.

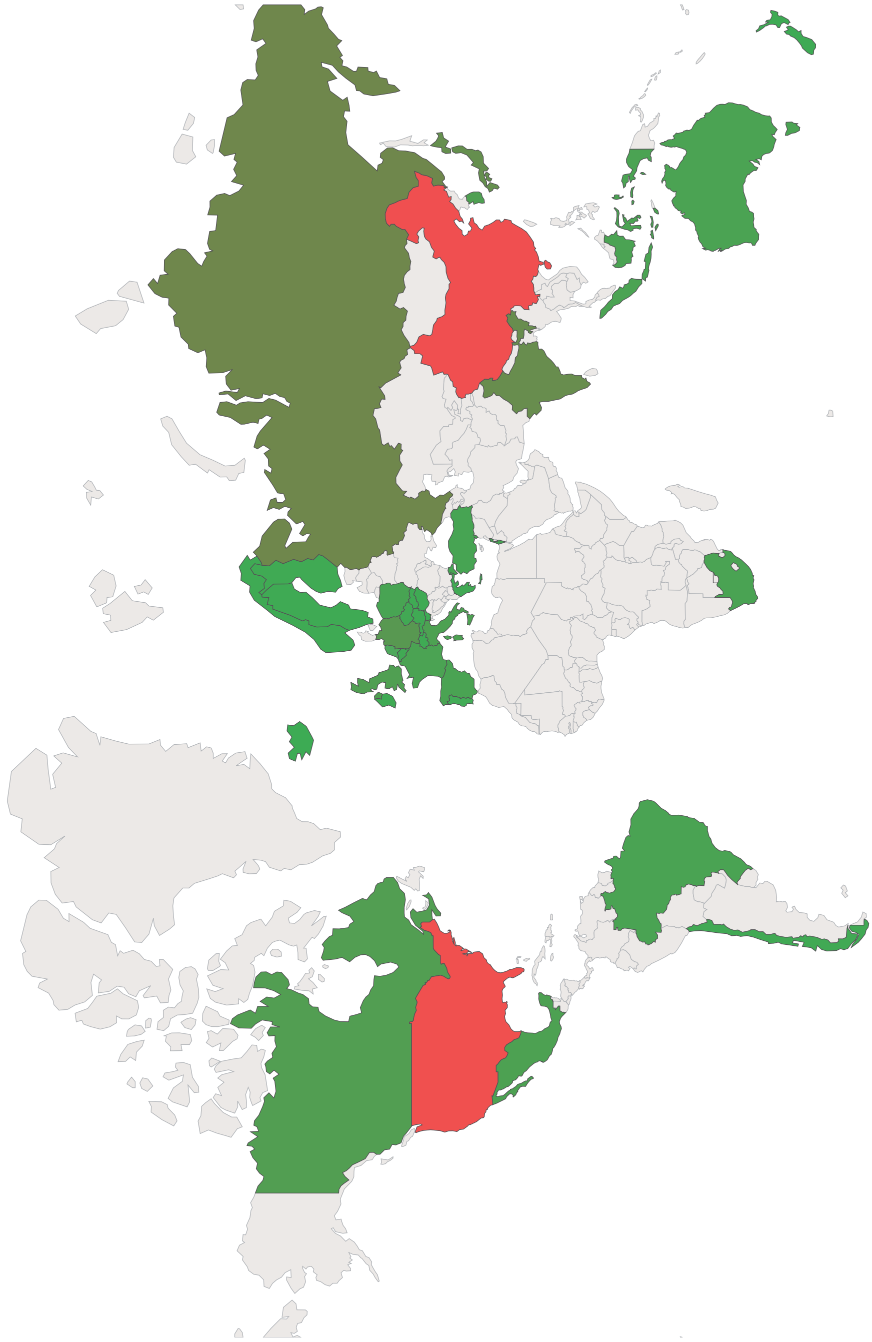
4.1 Background

4.1.1 Three Levels of Visual Communication

Ruedi Baur [117] states that the work of a graphic designer is to produce artifacts or documents (*e.g.,* websites, printed documents, or signage systems) that communicate with a user on three levels: an interpretation level, an orientation level, and an information level. On the *interpretation* level, the artifact or document should immediately reveal what it is, what its purpose is, and what it is about; it should answer the question "What is this?" On the *orientation* level, specific elements of the document should provide cues on how it should be read and used; these should answer the question "How does this work?" On the *information* level, the document should reveal the content and meaning of the communication; this should answer the question "What is being communicated, and what does it mean?"

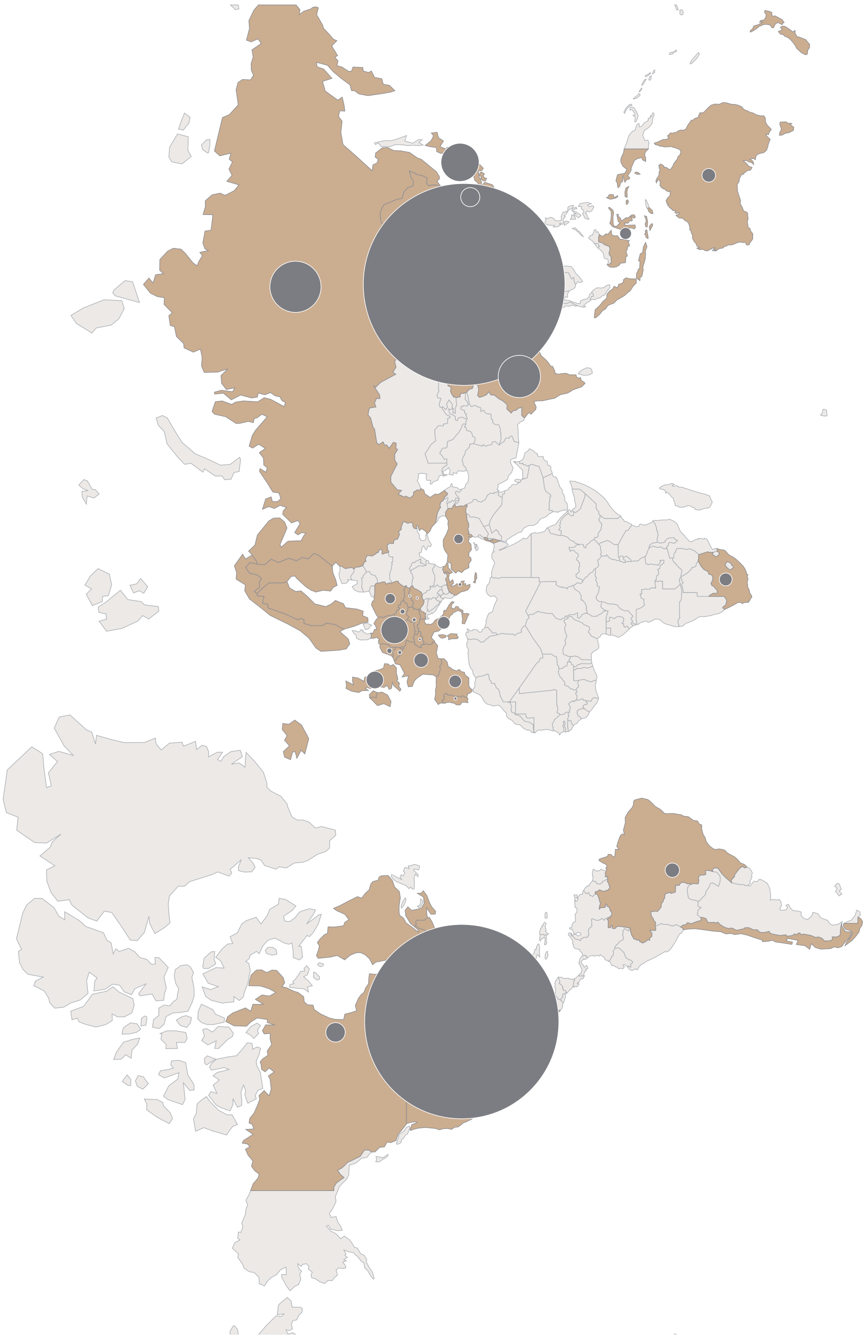
4.1.2 Pollution Maps?

A simple Google search of the terms "co2 pollution map" reveals that most maps that encode pollution related data use heatmap or choropleth encodings (see FIGURE 4.2). These generally fail on the interpretation level, as they do not communicate directly (*i.e.,* without the use of text) what the data are about, and because they could just as well represent population densities or temperature. In addition, most heatmaps that appear in the results of the Google search use spectral (or *rainbow*) color scales, which are known to have limitations [191]; and choropleth maps are generally ill-suited for representing density data like pollution [52], typically because 1) larger areas tend to provide lower density measures; 2) frequency fluctuations are 'erased' within a given area, which leads to a generalization effect; and 3) the actual location of area boundaries is often arbitrary (*e.g.,* administrative or enumeration zones) [189]. As an alternative, Langford and Unwin suggest using either *dasymetric mappings*



(A) A choropleth map of CO2 emissions.

FIGURE 4.2: Two 'standard' CO2 pollution maps (Figure is continued on next page).



(B) A bubble overlay map of CO_2 emissions.

[276] or *pseudo-3D density surfaces* [189]. Other possibilities include using bubble chart overlays or *cartograms* [152]. However, these all fail on the interpretation level, for the same reasons as mentioned above.

To produce a visualization that is effective on the interpretation level, we need to create a visual metaphor that allows users to immediately connect the visual features of the display with the semantic nature of the data.

4.1.3 Designing for Interpretation

Michael Danziger [147] declares that visualization designers "should try to identify design vectors that position and emphasize infovis as part of our intellectual, social, and cultural experience, rather than a cryptic, computational artifact that requires an explicit set of skills to decode. [...] Without an immediate cue as to the nature of the data being visualized, a non-committed viewer is likely to skip the graphic." He encourages the development of a semantic design approach for visualization, based on the creation of a structured visual language.

A noticeable, working example of this is Wattenberg and Viégas' much acclaimed *Wind Map* [53]. Interpreting this visualization is almost instantaneous, as the animated trails effectively convey the impression of wind blowing across the map. An amazingly rich source of inspiration for such semantic designs can be found in the film industry, and more precisely in animated films and special effects. Our design of the *CO2 Pollution Map* builds upon some of these techniques.

4.2 Designing CO2 Emissions

4.2.1 Metaphor

To design for interpretation, we need to find the right *visual metaphor*. Unlike most metaphors used in infovis, which serve to explicit structural components of the data, or to describe interaction techniques, this metaphor should convey semantic information about the data.

The first question to answer is: "What does the data's topic make me, and others, think of?" For the *CO2 Pollution Map,* we claim that the dark-gray smoke that emanates from exhaust pipes and chimneys reminds us of the presence of CO2 in the air. Therefore, using the combined potentials of d3.js [47], javascript, and SVG, we extend the bubble chart overlay encoding by transforming bubbles into animated smoke particles, and introduce a visual metaphor for *CO2 emissions.*

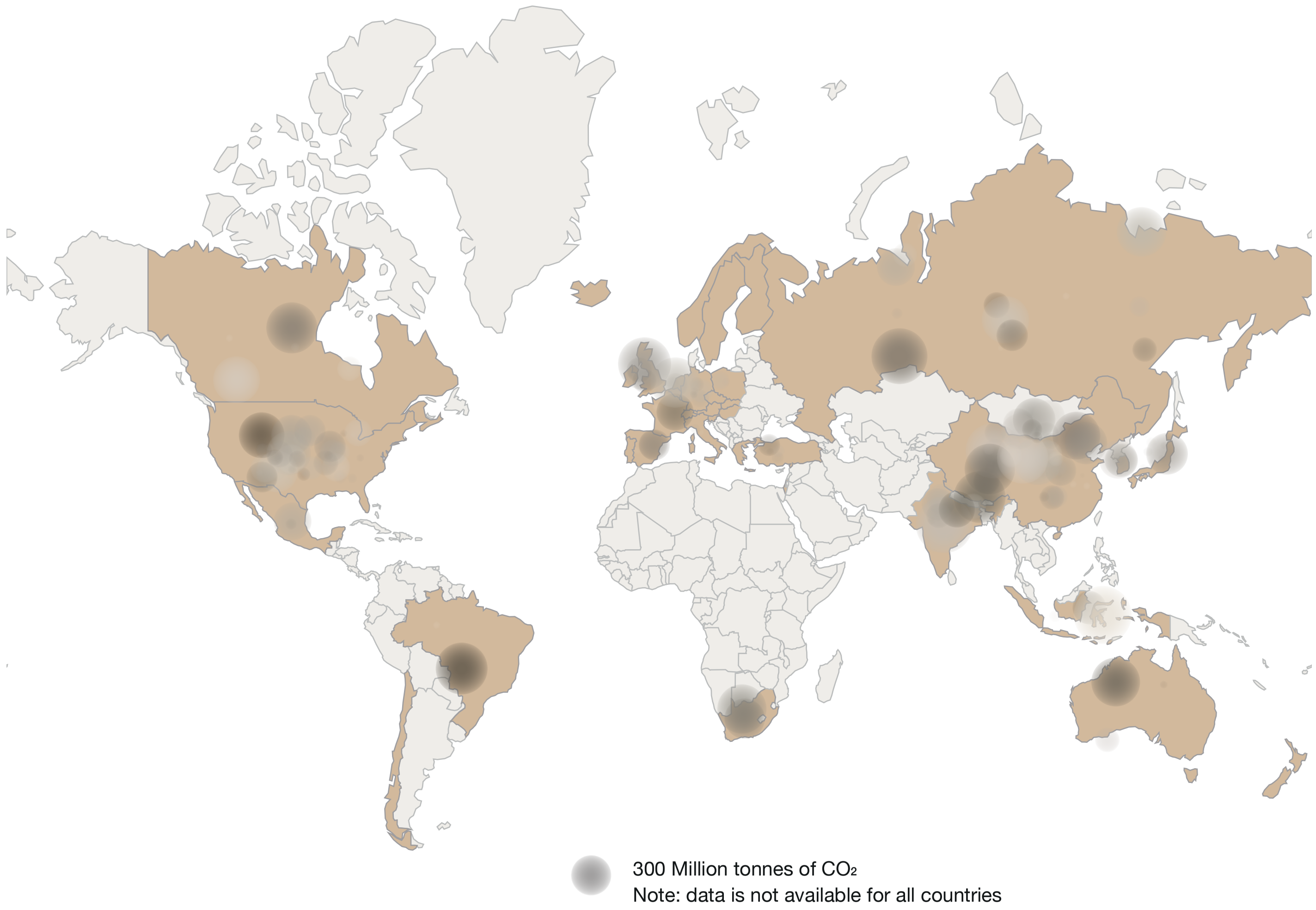
4.2.2 Particle system

In motion graphics and film post-production, a common way of designing smoke is to use a particle system [64]. *Particle systems* are a collection of graphic objects (*particles*) that are dynamic, chaotic, and generated from a source (*emitter*) [229]. Important attributes of a particle system are: 1) *emitter type,* 2) *particle type,* 3) *number* of particles emitted (nParticles) within a given *time period* (tEmit), 4) *evolution* of individual particles over life, and 5) *color* of particles. Note that we will not describe all the parameters of these attributes in detail. For more information, refer to [[66], [229]].

To prevent our particle system from conflicting with the mapping of the data, which relies on *position* (*i.e.,* the emitter) and *number* of visual marks (*i.e.,* nParticles) as *encoding* visual variables [183], we use five other free variables for the metaphor: *animation, shape, size, opacity,* and *color.*

For 1), we use a *box* emitter. In two-dimensional space, boxes are rectangular areas in which particles are generated at random X, Y positions. We use the SVG bounding boxes of countries (previously rendered using a *d3-geo-projection*) for which data are available. For 2), we use *feathered spheres,* as they are SVG compliant (*i.e.,* circles with a blur effect), and well-suited for designing smoke. For 3), we determine nParticles for each country by transforming the raw values into proportions of the maximum value (*i.e.,* a proportion of the highest amount of CO2 emissions), which we then map to a [0, 20] scale. We set the maximum number of particles to 20 due to technical limitations. SVG elements populate the *HTML Document Object Model* (DOM) tree, and the more nodes there are, the more computational power is required for rendering the tree. Thus, the highest of all CO2 emissions (≈6000 million tones in China in 2007) is displayed using 20 particles, each of which roughly encodes 300 million tones of CO2. At this point, we also set tEmit to 4 seconds (s). For 4), the life of each particle (tParticle) is set to match tEmit (*i.e.,* tParticle = tEmit = 4s). This way, when a particle *dies,* it is immediately replaced by a new one, insuring that at any point in time, nParticles are displayed for a given country. The evolution of particles over life is composed of three phases: *birth* (0s), *full growth* (tParticle/2 = 2s), and *death* (tParticle); and consists in the animation of two variables: size and opacity. At birth, particle size and opacity are set to 0. At full growth, particle size is set to 10 pixels, with a degree of randomness of [-5, 5] pixels, and opacity is set to 0.5. We use size randomness to ensure that particles never look exactly the same. This gives the smoke a more 'realistic' look. In addition, the 0.5 opacity nicely blends overlapping particles, also contributing to the realistic look. At death, particle size and opacity are both set back to 0, and 'dead' particles (tParticle 4s) are removed. Finally for 5), each particle is given a random color selected along a linear [dark gray, light gray] color scale.

FIGURE 4.3: In the Smog: the *CO2 Pollution Map*. Light gray countries indicate that data are unavailable.



4.3 Discussion

4.3.1 Benefits and limitations of the CO2 Pollution Map

The *CO2 Pollution Map* has the benefit of immediately conveying the semantic nature of the data. Informal feedback has confirmed that people interpret "pollution" from the smoke effect. In addition, the map helps uncover certain specificities of the data that other mappings cannot. While the number of visual marks encoding and the use of animation may be sub-par considering the desired preattentiveness of visual mappings, the overall density of smoke effectively reveals the concentration of CO2 in different regions. For example, individual European countries emit relatively low amounts of CO2 (compared to China or the USA), but the map shows that together they emit quite a lot (see FIGURE 4.3). Traditional density maps (*e.g.*, choropleth maps) cannot show this, as they do not reveal densities across country borders.

However, there are also limitations to this design. As it is, extracting specific values is almost impossible. Update is not instantaneous, and precise values are difficult to compare. The 300 million tones per particle limitation also forces a rough encoding of the data for several countries: Iceland, Finland, and Chile for instance emit less than 100 million tones of CO2 per year, meaning that no particle ever appears over them. A finer encoding could be achieved by re-implementing the visualization using HTML Canvas instead of SVG, because Canvas elements do not populate the DOM. Nevertheless, the use of animation and the randomness of particle positions within the emitter would still prevent precise readings. For example, FIGURE 4.3 shows a high concentration of smoke over the center of the USA, while it should be expected to be centered around the North East and West coasts. This is due to the aggregation of the data on a national scale (*i.e.*, to the density calculation), and to the box emitter we used. A solution would be to find non-aggregated data, *e.g.*, on a local scale, which would allow emitters to be changed to *points*, *i.e.*, invariant

X, Y positions centered on individual cities—as is done for dasymetric mappings [276]. However, similarly to the problems induced by dasymetric mappings (reported in [189]), this would increase the density of the visual marks, *i.e.,* nParticles, which would still make it almost impossible to extract precise values.

Overall, the purpose of this visualization is to deliver an *impression* of pollution concentrations. It is not designed to perform advanced visual analytic tasks. Tversky *et al.* [258] have indeed emitted warnings about the use of animation, due to the difficulty of perceiving and conceptualizing it. However, Lockyer *et al.* [194] have shown that motion textures can successfully communicate affect. We believe more work is needed in the vein of Huber & Healey's analysis of the perceptual properties of motion [178] to explore how this impression, and the potential affect it communicates, might alter users' ability to perform complex analytic tasks.

4.3.2 Early evaluation

We published the *CO2 Pollution Map* in the beginning of 2014 in a data-journalism-like article—*the CO2 Pollution Explorer* [8]—on the Mediapart Club website (a French online news outlet) [42], and on visualizing.org. It was referenced by visualisingdata. com as one of the "Best of the visualisation web... January 2014" [67]. By April 2014, the visualization had received over 3,000 individual browser connections, and the average connection time was ≈2min. These simple indicators suggest that the *CO2 Pollution Map* successfully captures viewers' attention. At this stage, we expect this success to be the result of the emphasis we put on designing for interpretation, *i.e.,* on designing the visual metaphor. Finally, note that we do not report more recent results here, as we have used the *CO2 Pollution Explorer* in another study, which will be presented in Chapter 6.

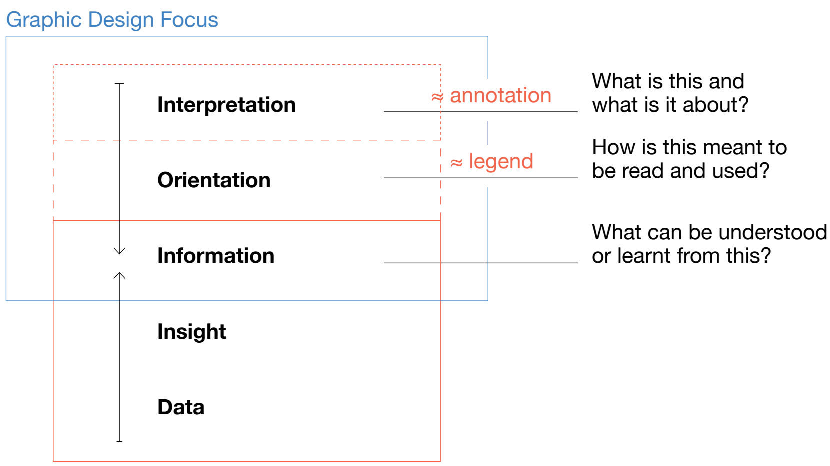


FIGURE 4.4: From interpretation to data, and back again: a framework for information visualization design that combines the top-down process assumed by graphic designers, and the bottom-up process traditionally assumed by infovis designers.

4.4 A Framework for Thinking about Visualization Design

From the design considerations that went into our implementation of the *CO2 Pollution Map*, we have derived the framework shown in FIGURE 4.4 to help rethink visualization design from a visual communication perspective. This framework is inspired by the three levels of visual communication described in Section 4.1.1, which show a clear hierarchy: 1) when a user encounters a new document or artifact, s/he must first interpret what it is and what it is about; 2) s/he must then be able to detect a certain number of cues that will help parse the document or artifact; and finally 3) s/he must be able to understand the information the document delivers. A simple analogy can be made with a book. First, a reader must detect that the 'object' is a book that holds information, and needs to rapidly determine what it might be about—this is generally the purpose of the book's physical design and of its cover. Then, the reader must detect how to read the book—this is generally the purpose of typographic hierarchy, tables of contents, *etc.* Finally, the reader should be able to read the information or the story contained in the book—this is dependent on the reader's literacy.

In our framework, we relate this first hierarchical approach to a *top-down* process for finding information. Users generally enter this process at the interpretation level, and make their way down to the information. Meanwhile, a simplification of the visual analytics process shows another hierarchical approach to finding information. This process starts with data, which users spend time exploring and analyzing to find interesting insights. These are then structured and transformed to create communicable information. This *bottom-up* process for finding information is the one traditionally assumed in information visualization design.

FIGURE 4.4 illustrates how these two processes can be combined to create a continuum. From a design perspective, it is important that visualizations start with data, and go all the way to the top of the framework, *i.e.*, to the interpretation level. Many communicative information visualizations already attempt to do so, by extending the bottom-up approach to the orientation level. They provide legends and other narrative cues [238] to suggest how the visualization should be 'read.' Some even push all the way to the interpretation level, but these tend to rely only on annotations or embellishments, which have several limitations. Annotations are not at all preattentive, and require viewers to switch between literacies, *i.e.*, between visualization and textual literacy; and embellishments are often very controversial, as mentioned in Section 2.2.2. Our design of the *CO2 Pollution Map* has shown that an alternative can be to use free visual variables to encode semantic information about the data.

Finally, it is important to consider that users are most likely to enter the framework at the top, *i.e.*, at the interpretation level, which is why we stress the importance of considering the **context-interpretation cost**. Once this cost is overcome, users should then make their way down to the information level. This requires they be visualization literate. However, we posit it is important they do not stop there. Visualizations should provide them with means for exploration, *i.e.*, with ways to go all the way down to the data level, only so that they can make their way back up again to the information level, but with their own questions and understanding of the data. Users may then loop through this cycle until a satisfactory amount of information is extracted.

4.5 Conclusion

This chapter has focused on the **context-interpretation cost**, and has addressed the following design problem:

Q2: How can visualizations be designed to help people interpret their context, *i.e.*, the semantic nature of the data they present?

I have presented the design of the *CO2 Pollution Map,* which I developed with Jean-Daniel Fekete. This design is inspired by the practices of graphic and motion design, and concentrates on helping people interpret the context of the visualization (or semantic nature of the data it presents) using free visual variables (**Q2**). An early evaluation has shown that the visualization successfully captures viewers' attention, and informal feedback has comforted us in the idea that the smoke metaphor we used conveys the idea of pollution. Our motivation was to propose a way to rethink the ways in which visualizations can communicate the topic of a dataset to a viewer. We have shown that the considerations that went into the implementation of the *CO2 Pollution Map* can be extended into a novel framework (FIGURE 4.4) for thinking about information visualization design. Overall, this work highlights the impact of considering the interpretation level on visualization design choices, and intends to assist designers rethink visualizations so that they help casual audiences overcome the **context-interpretation cost**.

Although the adage mentions it is important not to judge a book by its cover, many people do so when there are several on a bookshelf, and I believe that this is an important issue for many visualizations. The cover is what will draw a user into the document, by providing relevant contextual information related to the content. Thus, the purpose of the interpretation level of visual communication is to provide users with enough contextual information so that they can determine whether a document or artifact is relevant to their information seeking activities. However, providing context for an information visualization is not necessarily a matter of insight into what data are about. Free visual variables can also be used to communicate 'meta-data' about the visualization, like who the designer or publisher is. Graphic design is often concerned with creating *visual identities, i.e.,* graphical languages or systems that can be declined and applied to various documents, creating a consistent 'look and feel' for *e.g.,* a company (*e.g.,* Appendix I). Journalistic information visualizations can make use of this to provide users with a consistent experience across all visualizations of a specific online editor or publisher. Our framework should also be considered in this direction.

Finally, while we have adopted a design approach, we believe our framework holds the potential for interesting research into the way casual audiences engage with an information visualization, moving down from the interpretation level to the data level, and then looping between the data level and the information level. Ultimately, we hope that this work will serve as a foundation for further research into bridging the gap between information visualization and visual communication.

Chapter 5

Assisting Perceived Interactivity: Seeking Perceived Affordances for Infovis

In August 2012, Lars Grammel posted an article describing eleven online interactive visualizations that use different ideas and concepts of interaction design [1]. Grammel claims that "interactive data visualizations are an exciting way to engage and inform large audiences." He attributes this to the fact that they "facilitate a playful experience that is way more engaging than static infographics or video." But how do people know when a visualization is interactive or not? Although the importance of interaction is well recognized within the infovis community (see Section 1.1.2), and is often emphasized in the design of successful visualizations (see Section 2.4), until recently, most mainstream data graphics have been static (*e.g.*, infographics in news papers)—and many still are, even on the web. Thus, in a context like that of a data-journalism article, where visualizations are embedded with other media like text, it seems optimistic to assume that people know that they can or should interact with a visualization to find information; and as I have mentioned in the first two chapters of this thesis, this **perceived interactivity cost** can prevent casual audiences from engaging with the interactive potential of infovis, and therefore in the exploration of data.

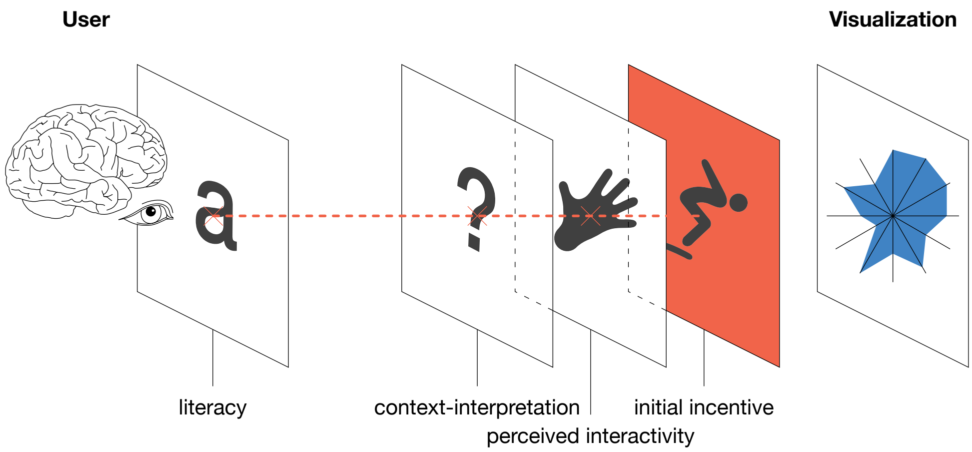


FIGURE 5.1: The **perceived interactivity cost**.

This chapter is based on a published paper entitled *Suggested Interactivity: Seeking Perceived Affordances for Information Visualization* [124], so any use of the term "we" refers to myself, Louis Eveillard, Françoise Detienne, and Jean-Daniel Fekete. It focuses on the **perceived interactivity cost** by addressing the following research question:

Q3: Do online users have a natural propensity to interact with visualizations—especially when these are embedded with text—and if not, how can we help these people detect the interactive potential of information visualizations?

To answer this, we first assess the need for means of suggesting the interactivity of visualizations to users, specifically when these are embedded in webpages with text. After that, we introduce the concept *Suggested Interactivity* (SI) and a design space for visual cues that can help users identify interactive graphical objects and areas on a webpage. We then propose several design considerations for applying these SI cues to charts embedded with text; we evaluate the benefits of three representative cues of our design space; and we draw some initial recommendations for design.

Information visualization designers often consider visualizations as isolated artifacts that users willingly come to view and interact with. However, many visualizations end up embedded in webpages with other media like text, and it is unclear whether users actually have a natural propensity to interact with these.

Most interactive features on the web use standard widgets (essentially buttons), which usually rely on metaphors of physical objects to suggest how they operate; they borrow affordances from their real world counterparts. These are not 'real' affordances in a 'Gibsonian' sense (see Section 2.2.3), since they do not support the physical actions of pointing, clicking, and (possibly) dragging with a mouse device; but they do suggest that an interaction is possible. Buttons, for example, designed with embossments and drop shadows (illustrating their mechanic origin), suggest that 'pressing' them is possible.

While effective, these analogies fall short when it comes to more abstract or symbolic interactive features, which mainly depend on design conventions. Hyperlinks, for example, use by default a specific visual variable, *i.e.,* color hue, and an additional visual mark, *i.e.,* an underline, to suggest that they can be clicked. This is 'heavy' design, as it requires two visual attributes to highlight a single difference with other textual elements, *i.e.,* interactivity; and the fact that a user knows that such highlighted text can be clicked on is purely conventional.

Interactive visualizations however, have neither convention, nor real world counterpart that can help suggest the fact that they can be interacted with—a pie chart does not afford eating! Therefore, if our initial prediction is true, *i.e.,* interacting with information visualizations embedded with text is not obvious to everyday Internet users, we ask: "how can we attract these users' attention to a visualization and suggest its interactivity through design?" To address our assumption, we first conducted three experiments on Amazon's Mechanical Turk (AMT) that confirm that a majority of people do not interact with visualizations embedded with text, even if these are more efficient for preforming given tasks. To address our main research question, we then surveyed 382 HTML5 and visualization websites to see how interaction designers make use of different visual cues to suggest the interactivity of *abstract graphical objects* and *areas*—where several interactive objects may be aggregated into a whole (*e.g.,* the bars in a bar chart). Based on this, we developed a design space for Suggested Interactivity. Finally, we conducted a follow-up study to evaluate the effectiveness of three SI cues applied to bar charts, which we believe are most representative of the diversity of our design space.

As such, the main contributions of this chapter are as follows:

* an assessment of the need for SI in cases where visualizations are embedded with text;
* a design space for SI;
* an evaluation of three different SI cues for bar charts, which we created using specific design consideration derived from our design space; which led us to
* initial recommendations for the design of SI cues for infovis.

Note that we adopt both an evaluation and a design approach here, as our immediate motivations for this work are to assess whether Suggested Interactivity is needed, and to help designers create cues that can suggest the interactivity of abstract features like information visualizations in a webpage.

This chapter assumes a non-conventional structure, and is organized in the following way. Since our assumption that people are unlikely to interact with visualizations when these are embedded with text is quite strong, and that to the best of our knowledge this has not yet been assessed, we begin by presenting our initial experiments that show the need for SI. After that, we provide a higher-level understanding of the possible reasons why online users lack interaction propensity; we extend Section 2.2.3 by introducing the concept of feedforward; we describe several new graphic standards for interface design, and discuss how these consider suggesting the interactivity of specific features; and we present previous work on the use of motion and icons in interface design to attract users' attention and to convey meaning. In Section 5.3, we then describe our survey of existing SI cues, introduce our design space, and provide a set of design considerations, which we use for creating three SI cues that we apply to bar charts. In Section 5.4, we present the design and evaluation of these cues, and discuss the implications of our early results for future designs. Finally in Section 5.5, we discuss possible extensions of our work for other visualization applications.

5.1 Testing Interaction Propensity

To verify whether online users are naturally inclined to interact with visualizations embedded with text, and to assess the need for SI, we conducted three initial experiments on AMT. For each, we used a series of seven simulated Wikipedia articles we created, which included both visualizations (bar charts) and text. We used data and text from the OECD's *Better Life Index website* [93], and grouped similar-topic indicators into specific articles (as is done on the OECD's website); topics were *Housing, Income, Education, Environment, Health, Safety,* and *Work-Life balance.* We reduced the original text for each article to limit the amount of contextual information it provided, in order to create a better information-balance between the text and the visualizations. We created specific sections in the articles for each indicator (either two or three, depending on the topic), and we displayed the bar charts next to the corresponding paragraphs; this design follows the traditional Wikipedia layout, with text to the left, and images (in this case charts) to the right (see FIGURE 5.2).

We chose to simulate Wikipedia articles for ecological validity, since Wikipedia will undoubtedly soon provide tools for building interactive visualizations—the markup already supports the creation of static charts (*e.g.,* [50]). While it is arguable that this choice may bias users' propensity to interact with visualizations (since there are few in contemporary Wikipedia articles), it is a realistic setting. In the same line, online news articles like those of the Guardian, which heavily rely on traditional media like text, now integrate more and more interactive graphics. Thus, we believe Wikipedia is a good and timely environment for testing people's propensity to interact with visualizations.

To save screen-real-estate, we limited the default labeling of the charts to OECD averages, and highlighted the corresponding bars. As there were either two or three charts per article that presented data about the same OECD countries, we implemented a simple *brushing-linking* technique to highlight identical countries in different charts simultaneously;

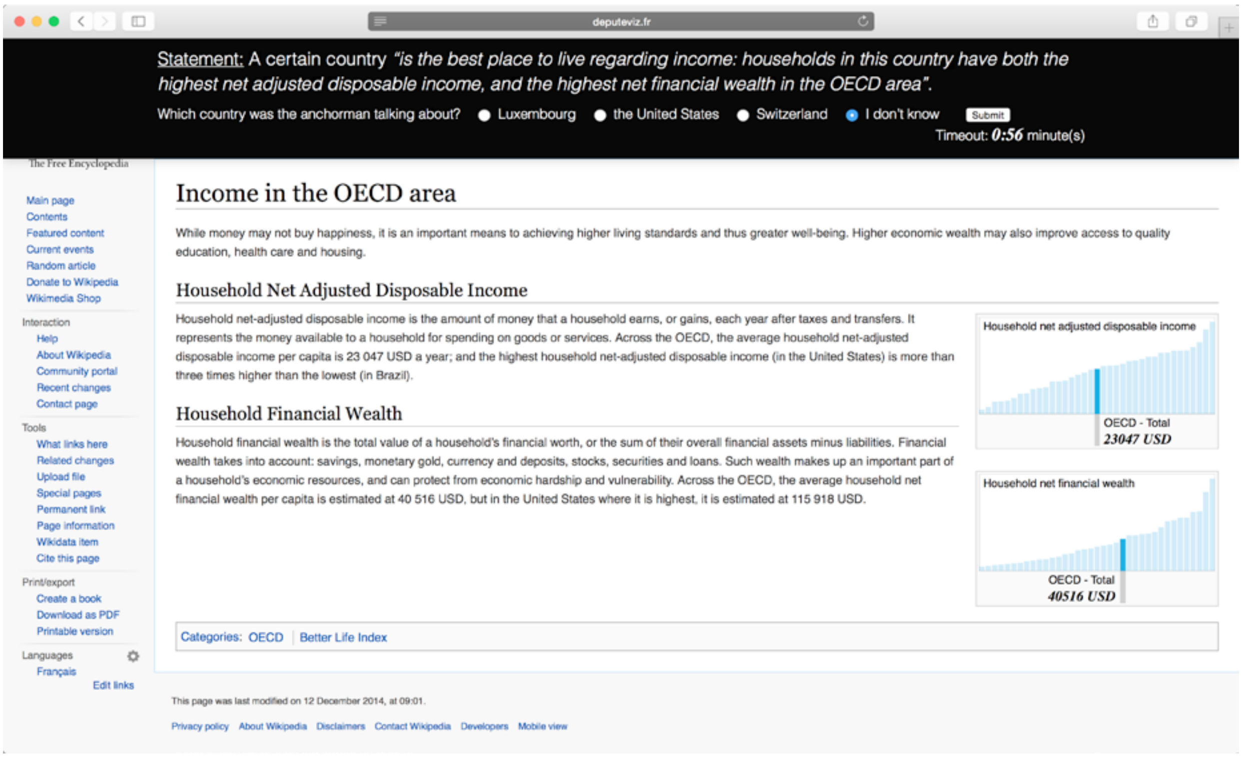


FIGURE 5.2: Screen capture of article on Income.

this also displayed their labels and precise values. Thus, interaction was necessary to extract specific values from the charts, and this simple hover interaction was what we expected participants to discover.

We created a simple fact-checking task for each article, as fact-checking is a common activity on the web; this was also done to encourage people to go straight to the point, *i.e.*, to find the specific datapoint we asked them to target, without needing to seek for contextual information. Each article corresponded to a trial, resulting in seven trials per participant, the order of which we simply randomized to prevent carry-over effects.

Our first experiment was conducted to assess whether people are inclined to interact with charts to carry out fact-checking tasks. Our second experiment was conducted to make sure that the charts were indeed more efficient than the text for performing the tasks we created, and that if given the chance to discover their interactivity, participants would be more inclined to use them. Finally, our third experiment was conducted to make sure our choice of simulating Wikipedia articles did not bias participants' propensity to interact with the charts.

5.1.1 Experiment 1

5.1.1.1 Design

Scenario—To situate the fact-checking activity, we provided participants with the following simulated task scenario:

> This local news anchorman [a fake photo of an anchorman was displayed above the scenario] recently made a series of statements on living in different countries in the OECD area. We recorded seven of these statements, but unfortunately, we did not record the countries he was talking about. Luckily, we have come across a series of Wikipedia articles that discuss the topics of each of the anchorman's statements, and we have narrowed down the number of possible countries to three for each statement. Please help us find the corresponding countries for each statement in this series of Wikipedia articles. Note that we will require you to do this as fast as you can, as we have set a specific timeout for each article.

Tasks—We designed the fact-checking tasks specifically to make the visualizations more efficient than the text for retrieving the necessary information (provided that participants interacted with them—see Appendix J). For each article (or trial), we set a multiple-choice extraction task with 3 possible answers and an "I don't know" option, which required participants to consider all the indicators presented at once; a typical example for an article that presents three indicators would be: "Which country has the highest rate of X, the lowest rate of Y, and an average rate of Z." Each possible answer was hand-picked to complicate reliance on potential background knowledge. For example, the possible answers for the article on Income, where the task was to find the country in which households have both the highest income and financial wealth, were Luxembourg, Switzerland, and the United States, *i.e.*, three countries where one might expect that income is high, but does not necessarily know in which it is the highest (see FIGURE 5.2). To find the correct answer, participants simply needed to brush over one of the visualizations until they found a country that met the first requirement, and then check its performance across the other indicators. We also purposefully chose combinations of min/max tasks to make the questions as highly-congruent as possible (see Section 3.2.3). However, to balance out the study design, and to make it possible to perform the tasks using the text alone, we added specific mentions of the possible answers in the paragraphs corresponding to each indicator—this created a certain redundancy between text and charts. If participants were to use the text, they would need to read through all paragraphs and memorize how well each possible country performs across all indicators.

Procedure—Upon accepting the HIT on AMT, participants were directed to an external page for the study. On this page, they were first asked to complete a pre-study to make sure that they had the necessary English skills to participate in our experiment, and that they were willing to comply with instructions—the pre-study was an intermediate English reading comprehension test taken from [51]. Participants who failed the pre-study were not allowed to continue. Those who succeeded were then asked to fill out a short, anonymous demographic survey, were given the scenario, and were administered the study. Before each trial, participants were instructed the task, and were invited to click on a "Display Wikipedia page" button to display the article. As mentioned in the scenario, we set a time-out for each trial, based on the number of words in the article; we divided this number by 200, in accord with an average reader's *word-per-minute* (wpm) *speed-reading score* [[32], [133]], and rounded the result down to the nearest half-minute, in order to force participants to be quick. Clicking on the "Display Wikipedia page" triggered the countdown. Finally, at the end, participants were asked to fill out another shot survey about the study.

Hypotheses—We had two simple hypotheses for this experiment:

* **H1.1:** a majority of participants will not know that the charts are interactive, and therefore they will not use them to complete the trials; and
* **H1.2:** a majority of participants who 'discover' the interactivity of the charts will use them throughout all subsequent trials, as they are in principle more efficient.

Participants—We recruited 70 participants on AMT who were required to have a 98% acceptance rate and a total of 1000 or more HITS approved. We removed the work of 2 from the collected data, as these had taken the HIT on a mobile device—such devices do not support brush interactions, *i.e.*, hover interactions—but we paid them nonetheless. This resulted in a subset of 68 participants, all of which were native English speakers.

Coding—We traced participants' low-level activity on the external page using a custom built system, and we first counted the number of *brush* (hover) interactions each participant performed. However, since such indicators are often noisy, we also counted what we refer to as *decisive brushes, i.e.,* brush interactions over bars related to the answers participants gave that lasted more than 250 ms (so that participants had time to see their effect on the display). For example, if a participant answered "the United States" to the question for the article on Income (mentioned above), we coded every brush interaction that lasted longer than 250 ms over a bar encoding US data as a decisive brush. In addition, since we did not reset the display to its original state after user interactions (with the OECD average highlighted and labeled), we also counted one decisive brush if the last bar to be highlighted of a series of brush interactions that each lasted less than 250 ms was related to the answer participants gave. Note that in this chapter, when we claim participants *used* the charts to find answers, we mean they performed at least one decisive brush. We then counted both the number of trials in which each participant performed brush interactions and decisive brushes, and for each we coded 1 when such interactions where performed in all subsequent trials to the one in which the interactivity of the charts was discovered, and 0 otherwise. Finally, we coded participants' answers 1 when correct, 0 when the "I don't know" option was submitted, and -1 when incorrect.

5.1.1.2 Results

All the analyses and discussions in this chapter are based on estimation, *i.e.*, point estimates and effect sizes with confidence intervals (95% CI), with respect to the concerns and recommendations in [[110], [145], [153]]. Point estimates and 95% CI are based on 10,000 percentile bootstrap replicates of the statistic (in this case percentages and means) applied to the data [130]. The proportion 95% CI are calculated using the VassarStats.net web-application, which nicely documents how this is done.

We first inspected participants' scores. As the questions and tasks were overall quite simple, and as the default answer was "I don't know" (= 0), we removed the work of all participants whose total score was below or equal to 0 from further analysis—we considered these to be either random clickers, or people who only provided answers based on *a priori.* This resulted in a subset of 59 participants.

42.4%, 95% CI [30.6%, 55%] of these participants (25/59) performed at least one brush interaction, and **28.8%**, 95% CI [18.8%, 41.4%] (17/59) performed at least one decisive brush. Thus, **68%**, 95% CI [48.4%, 84.3%] of participants (17/25) who performed a brush interaction also performed a decisive brushes.

52%, 95% CI [33.5%, 69.9%] of participants (13/25) who performed brush interactions performed at least one in all seven trials, and **60%**, 95% CI [40.7%, 76.6%] (15/25) performed at least one brush in every subsequent trial to the one in which they discovered the interactivity of the charts.

Finally, **58.8%**, 95% CI [36%, 78.4%] of participants (10/17) who performed decisive brushes performed at least one in all seven trials, and **88.2%**, 95% CI [65.7%, 96.7%] (15/17) performed at least one in every subsequent trial to the one in which they first performed a decisive brush.

5.1.1.3 *Discussion*

Although admittedly we had expected that less participants would discover the interactivity of the charts, our results still confirm **H1.1.** We suspect this higher number may be due to the layout of the Wikipedia articles, as the charts were set to the right hand side of the page, next to the scroll bar. Participants whose displays were too small to show the whole webpage may have hovered over the charts while moving their cursor to scroll. This is interesting though, as it could suggest that the layout itself can be used to help people discover interactive content.

H1.2 is also confirmed, as a majority of participants who discovered the interactivity of the charts continued to brush them throughout all subsequent trials; and this is particularly true for participants who performed decisive brushes. This seems to indicate that participants who discovered the charts' interactivity and understood how to use them perceived the charts as more efficient than the text.

5.1.2 Experiment 2

5.1.2.1 Design

To extend our confirmation of **H1.2**, and to ensure that our task-design did indeed make the visualizations more efficient for extracting the necessary information, we conducted a second experiment on AMT. The design of this experiment was identical to that of the previous, with the exception that in trials 3, 4, and 5 (out of seven, and whatever the article) we removed all the textual information, and made the visualizations much larger, laid them out on the left hand side of the screen, and explicitly mentioned that they were interactive (see Appendix K). This was done to force participants to use the charts. Thus, this experiment consisted of two initial trials in which the charts were embedded with text, followed by three trials in which there was no text, and completed with to final trials in which the charts were once again embedded with text; the scenario, tasks, and procedure were kept the same as before.

Hypotheses—We had four hypotheses for this experiment:

* **H2.1:** all participants will interact with the charts in trials 3, 4, and 5;
* **H2.2:** a majority of participants will use the charts in the last two trials;
* **H2.3:** there will be good evidence that more participants interact with the charts in the last two trials than in the first two; and
* **H2.4:** as more participants should interact with the charts in the last two trials than in the firs two, and as the charts are hypothetically more efficient, there will be good evidence that participants complete the last two trials faster than the first two.

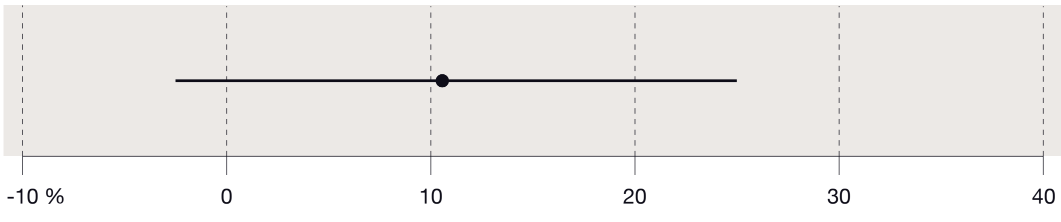
Participants—We recruited 70 different participants on AMT, in order to make sure they would not be biased by the first experiment. However, this time we only retained the work of 47 participants whose total score was higher than 0; all were native English speakers. We then coded the data in the same way as in Experiment 1.

5.1.2.2 *Results*

We first inspected the number participants who interacted with the charts in the first two trials. In trial 1, **12.7%**, 95% CI [5.9%, 26.4%] (6/47) performed at least one brush interaction, and **6.4%**, 95% CI [2.2%, 18.6%] (3/47) performed at least one decisive brush. In trial 2, **8.5%**, 95% CI [3.4%, 21.3%] (4/47) performed at least one brush interaction, and **2.1%**, 95% CI [0.4%, 12.7%] (1/47) performed at least one decisive brush.

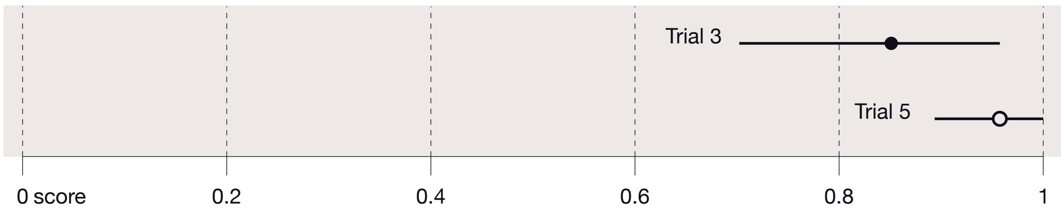
We then inspected the number of participants who interacted with the charts in the trials in which only the charts were displayed. In trial 3, **80.5%**, 95% CI [72.3%, 92.6%] (40/47) performed brush interactions, and **72.3%**, 95% CI [58.2%, 83.1%] (34/47) performed decisive brushes. In trial 4, **95.7%**, 95% CI [85.7%, 98.8%] (45/47) performed brush interactions, and **78.7%**, 95% CI [65.1%, 88%] (37/47) performed decisive brushes. In trial 5, **95.7%**, 95% CI [85.7%, 98.8%] (45/47) performed brush interactions, and **91.5%**, 95% CI [80%, 96.6%] (43/47) performed decisive brushes.

Out of the seven participants who did not interact with the charts in trial 3, four discovered the interactivity in trial 4 and continued to interact with the charts in trial 5; one discovered the interactivity in trial 5; one discovered the interactivity in trial 4, but oddly did not continue to interact in trial 5; and one simply never interacted. We checked what these last two participants did in detail, to make sure they were not random clickers. The first seemed to have made a lucky guess in trial 3—or she already knew the answer—then she attempted to interact with the charts in trial 4 to find the answer, but it seems she failed since she submitted the “I don’t know” option; she then directly submitted the “I don’t know” option in trial 5. The second simply submitted the “I don’t know” option for all



(A) Difference between proportions of participants who performed decisive brushes in trials 3 and 5 (within participants who performed hover interactions) with 95% CI.

(B) Mean time spent completing trials 3 and 5 with 95% CI.

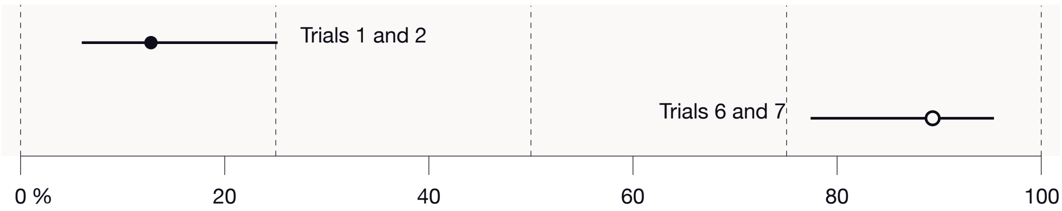


(C) Mean scores for trials 3 and 5 with 95% CI.

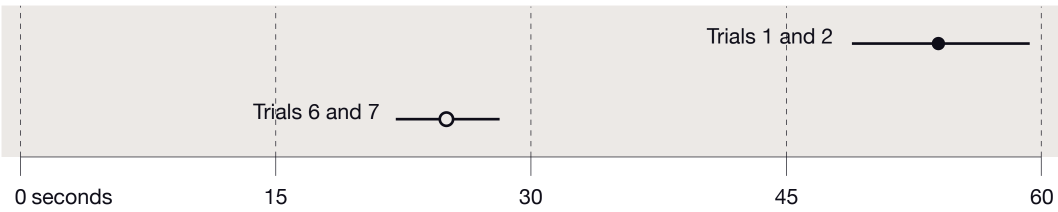
FIGURE 5.3: Participants' progression through trials 3 to 5.

three trials. Interestingly however, this participant did interact with the charts in trials 6 and 7, but she performed no decisive brushes, meaning she based her judgement on the text or on background knowledge.

These two cases made us wonder whether there may have been a visualization literacy problem (see Chapter 3) [122], since participants not only had to interact with the charts to find the answers, they also had to know how to search visually for them. To check for this, we inspected whether participants showed signs of *progress* through trials [3–5]. To es-



(A) Percentages of unique participants who hovered at least once over one of the charts in trials 1 and 2, and in trials 6 and 7 with 95% CI.



(B) Mean time spent completing trials [1, 2] and trials [6, 7] with 95% CI.

FIGURE 5.4: Comparisons of aggregated results for trials [1, 2] and [6, 7].

timate this, we isolated participants who performed brush interactions, and calculated the proportions of these who performed decisive brushes in trials [3, 5]; we then calculated the difference between proportions (FIGURE 5.3 (A)). We also compared the mean time participants spent in trials 3 and 5 (FIGURE 5.3 (B)), and their mean scores (FIGURE 5.3 (C)).

After that, we inspected the number of participants who interacted with the charts in the last two trials. In trial 6, **82.9%**, 95% CI [69.9%, 91.1%] (39/47) performed brush interactions, and **70.2%**, 95% CI [56%, 81.3%] (33/47) performed decisive brushes. In trial 7, **89.4%**, 95% CI [77.4%, 95.4%] (42/47) performed brush interactions, and **72.3%**, 95% CI [58.4%, 83.1%] (34/47) performed decisive brushes.

Finally, we aggregated the results for trials [1, 2] and for trials [6, 7], in order to compare what participants did before and after they were 'forced' to use the charts; we inspected the number of unique users who interacted with the charts (FIGURE 5.4 (A)), and the mean time they spent completing the trials (FIGURE 5.4 (B)).

5.1.2.3 Discussion

Our results do not support **H2.1.** An important majority of participants did interact with the charts in trials 3, 4 and 5, but it seems they needed to progressively get 'used' to them. FIGURE 5.3 (A) shows some evidence of an increase in decisive brushes between trials 3 and 5, as the lower boundary of the 95% CI is only slightly below 0. This suggests that participants needed the three trials to elaborate strategies for finding the answers in the charts. Similarly, FIGURE 5.3 (B) shows good evidence of a reduction of time spent completing the trials, which suggests that participants progressively became more efficient in finding the answers in the charts, and perfected their search strategies. We interpret this progress as a possible indicator for initial low visualization literacy. As fewer participants performed hover interactions in trial 3, we hypothesize that they may have initially preferred to avoid the charts, not necessarily because of lack of propensity to interact, but because of lack of strategies for finding the answers in charts. If a person lacks visualization literacy, the cost of interacting with a chart will be perceived as greater than the benefit, because the benefit is unknown. However, it seems that only three trials sufficed for most participants to overcome this problem, which indicates that learning to answer highly-congruent questions using bar charts can be done quite rapidly.

While this could have been a typical application case for the Bar Charts VL test described in Chapter 3 [122]—to make sure participants had the appropriate skills—we chose not to use it, as we feared it might bias participants' behavior, since it would prime them with visualizations, implicitly suggesting that there may be something to do with the charts.

H2.2, **H2.3**, and **H2.4** however, are all confirmed. This suggests that the charts are indeed more efficient, and that given the chance to discover the interactivity (and to elaborate effective search strategies), most participants will keep using the charts instead of going back to using the text. Thus, we consider that participants' lack of propensity to interact with the charts is not due to an efficiency problem; and while perceived efficiency may be an initial concern (due to low visualization literacy), it is rapidly overcome.

5.1.3 Experiment 3

5.1.3.1 Design

Finally, to ensure our results were not biased by the Wikipedia template, we conducted a third experiment which replicated Experiment 1, but for which we removed all Wikipedia styling attributes from the articles (see Appendix L). This was done to done to check that our results were not confounded by expectations participants may have had from Wikipedia articles. We kept the same overall layout, scenario (in which we replaced all occurrences of the term "Wikipedia" by "article"), tasks, and procedure.

Hypothesis—We had one simple hypothesis for this study:

* **H3:** results will be consistent with those of Experiment1, meaning that the Wikipedia styling did not bias participants' propensity to interact with the charts.

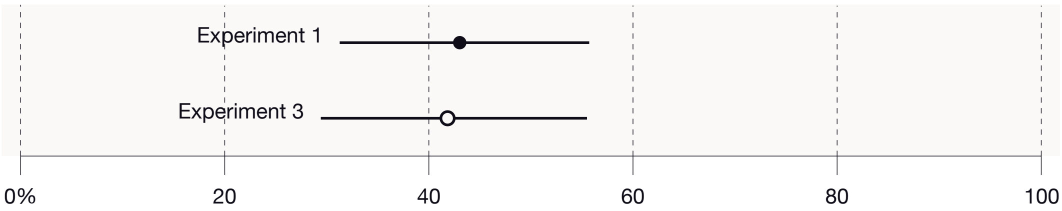
Participants—Once again, we recruited 70 different participants on AMT, in order to be able to establish comparison with the results of Experiment 1 in a between subjects design. We retained the work of 51 participants whose total score was higher than 0; all were native English speakers. We then coded the data in the same way as before.

5.1.3.2 Results

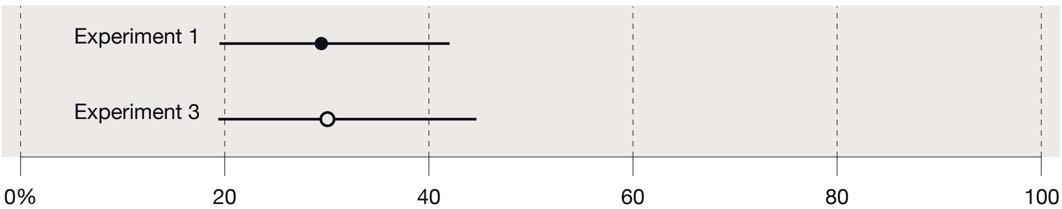
Our analysis was exactly the same as for Experiment 1; all results are shown in FIGURE 5.5, and are compared with those of Experiment 1 in a between subjects fashion.

5.1.3.3 Discussion

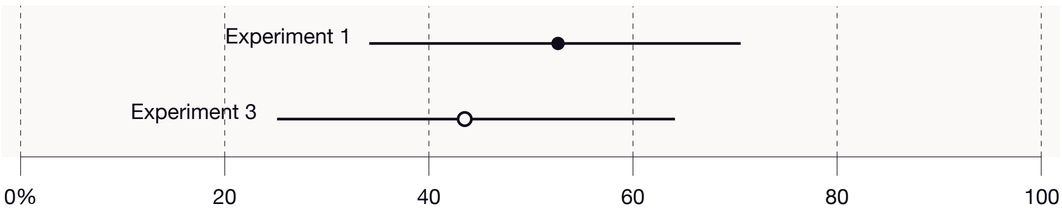
Although the 95% CI in FIGURE 5.5 are quite wide—which seems normal as



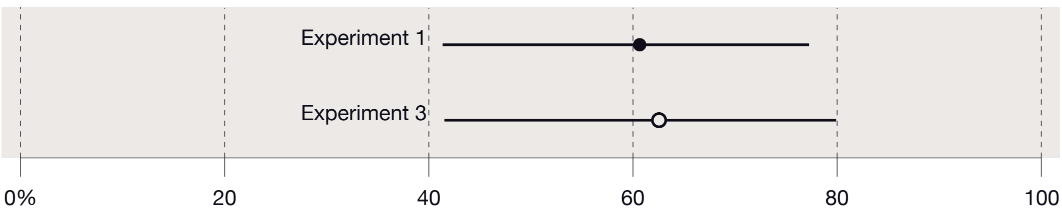
(A) Percentages of participants who hovered at least once over one of the chart in Experiments 1 and 3 with 95% CI.



(B) Percentages of participants who preformed at least one decisive hover in Experiments 1 and 3 with 95% CI.

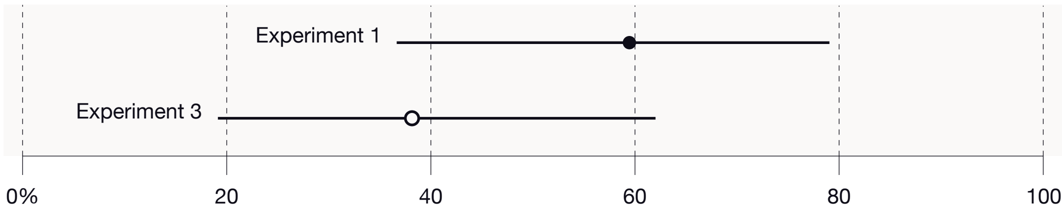


(C) Percentages of participants who preformed at least one hover interaction in all seven trials in Experiments 1 and 3 with 95% CI.

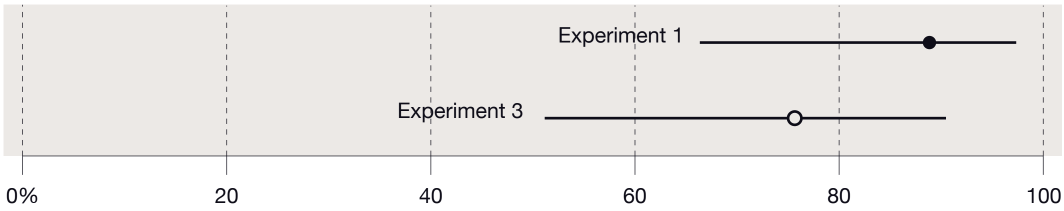


(D) Percentages of participants who preformed at least one hover interaction in every subsequent trial to the one in which they first performed a hover interaction in Experiments 1 and 3 with 95% CI.

FIGURE 5.5: Between subjects comparisons of results in Experiments 1 and 3 (Figure is continued on next page).



(E) Percentages of participants who preformed decisive brushes in all seven trials in Experiments 1 and 3 with 95% CI.



(F) Percentages of participants who preformed decisive brushes in every subsequent trial to the one in which they first performed a decisive interaction in Experiments 1 and 3 with 95% CI.

the estimation is based on only one bit of information, *i.e.*, percentages—results show no real evidence of a difference between Experiments 1 and 3. On the contrary, they show a high similarity (with a slight exception shown in FIGURE 5.5 (E)). This confirms **H3**, and suggests that the Wikipedia template did not bias participants' propensity to interact with the charts.

5.1.4 Initial Experiments' Discussion

Overall, our results show that a majority of people lack initial propensity for interacting with charts when these are embedded with text, whatever the styling of the webpage. This may in part be due to visualization literacy problems, but if people discover the interactivity, they are likely to rapidly learn how to perform the necessary visual queries for finding answers in charts. They will then continue to interact with them, as the charts are indeed more efficient for fact-checking tasks like the ones we created.

This indicates that people can be motivated to interact with visualizations if they are shown the possibility, and highlights the need for suggested interactivity (at least when visualizations are embedded with text).

5.2 Background

Finding ways to suggest the interactivity of graphical objects is an old issue for interface design. However, with the development of new graphic standards, especially on the web where designers and artists can easily create any new kind of interactive visuals, the concern is becoming of importance once again. In addition, while Segel & Heer have pointed out the use of *markers of interactivity* in certain narrative visualizations [238], this issue has hardly been considered by the infovis community—and as our initial experiments have shown, it needs to be address.

5.2.1 Passive Interaction

Information visualizations on the web often end up embedded with other media like text. Common examples of this can be found in data journalism articles like those of the Guardian [98]. Most of the time, these other media are static and do not suggest any interaction.

Passive interaction defines the changing or enhancement of a user's mental model while interacting with a system without modifying the system's model [[164], [245]]. In simpler terms, passive interaction occurs when reading a text, looking at an image or a visualization, or more generally, when receiving, decoding, and interpreting a new piece of information, without having to manipulate the medium—it requires little or no input device manipulation.

Passive interaction is a main component of what Bret Victor names *Information software* [71]. According to his definition, such software "serves the human urge to learn;" it initiates cognitive processes like learning, comparing, and decision making. He opposes this to *Manipulation software,* which "serves the human urge to create;" it helps a user construct and manipulate an external model: that of a system. Most visualizations on the web are intended to show viewers something new: they serve the urge to learn. However, they are not necessarily static, and users can manipulate them to display data in different ways. Thus, while Bret Victor speaks of software, we posit the same goes for online media. People have certain expectations from media, and if the main medium requires no interaction (*e.g.*, text), it seems unlikely that they will seek to determine whether other components of a webpage are 'manipulable.'

5.2.2 Feedforward

Section 2.2.3 has presented the concepts of affordances and perceived affordances; these are properties of an object that induce action in a user, *i.e.*, they suggest what is doable with that object. In parallel to affordances, Vermeulen *et al.* have highlighted the importance of providing *feedforward* in interaction design [264]; it "tells users what the result of their action will be."

However, while important and effective, perceived affordances and feedforward are not easily designed, especially when new interface design standards lean towards more abstraction. This is typically the case for visualizations, as they are abstract representations by nature.

5.2.3 New Interface Design Standards

Recent developments of platform-centric user-interface guidelines have approached the issue of perceived affordances in greater detail than before. Historically, Apple was one of the first companies to provide such guidelines—the *Human Interface Guidelines* (HIG)—in order to help third parties create consistent experiences when developing for their platforms. These guidelines explained which perceived affordances to use, and when to use them. For example, iOS 6's HIG emphasized the use of embossed, skeuomorphic cues to suggest that buttons are actionable. However, iOS 7 and 8's HIG propose a radical reorientation: they suggest designers "embrace borderless buttons," and mention that the use of "context, color and a call-to-action title [is enough] to indicate interactivity" [35].

Similarly, Microsoft's *Visual Studio 6.0* instructions for "creating a user interface" [102] mentioned that "a user interface also makes use of affordances. For instances, the three-dimensional effects used on command buttons make them look like they are meant to be pushed. If you were to design a command button with a flat border, you would lose this affordance and it wouldn't be clear to the user that it is a command button." However, ten years later Microsoft reverted its course and presented the *Metro Design* standard which states that "while lack of affordance and discoverability may sound discouraging, note that drop-down menus and context menus—other mechanisms for initiating actions—suffer similar problems" [68].

More recently, Google published its own set of guidelines for *Material Design*, which also describes a variety of ways to suggest interaction [23]. Depending on the importance of the element and its associated action, it can be made into a colored *floating action button*, a *raised button* or a *flat button*—the number of cues is proportional to the importance of its function. A fundamental aspect of *Material Design* is the use of motion to convey interactivity: "Perceiving an object's tangible form helps us understand how to manipulate it. Observing an object's motion tells us whether it is light or heavy, flexible or rigid, small or large. Motion in the world of material design is not only beautiful, it builds meaning about the spatial relationships, functionality, and intention of the system" [63]. Rather than copying the textures and shadows of physical objects, applying realistic motion to graphical objects helps with the perception of affordance.

5.2.4 Attracting Attention and Conveying Meaning

5.2.4.1 Attracting Attention with Motion

Motion is known to have major psychophysical benefits for capturing attention. As humans, our visual systems are extremely sensitive to fragments of natural motion [208]; our peripheral vision is also highly receptive to movement, and operates at a much lower resolution than our fovea [168]. We are able to view and interpret several motions simultaneously [227], and we can perform complex grouping tasks of moving elements. Research in visual science has shown that we perceive *causality* and *anima-*

cy in motion, primarily through perceptual, and possibly through modular processes [237]. This means we attribute *life, intentionality,* and *behavior* to moving objects without impeding higher-level cognitive processes or judgements. However, it is important to note that the graphical representation of motion (*i.e., animation*) is not always easy to perceive accurately or to conceptualize when it conveys more abstract meaning [257].

5.2.4.2 Conveying Meaning with Icons

The use of pictographic symbols to convey meaning is a classic of interface design, and more broadly of machine display design [[154], [195]]; these are commonly referred to as *icons.* Although there was originally some debate about their effectiveness [[112], [165], [200]], today icons are a major part of any interface, whatever the system or device, and have undoubtably greatly contributed to the success of personal computers. Huang *et al.* claim that "icons offer the perception of affordance, which can facilitate human-machine interaction in terms of ecological perception" [177].

Lodding [195] has proposed a taxonomy of icons for user interfaces, which is composed of three dimensions: representational, abstract, and arbitrary. *Representational* icons are "typical" images that serve as examples for general classes of objects. *Abstract* icons represent concepts, and use illustrations of real objects to refer to abstract ideas. Finally, *arbitrary* icons do not relate to any object in the world; they are "invented" and assigned a conventional meaning.

5.2.4.3 Combining Motion and Icons

The combination of motion (or animation) and icons can help clarify meaning, explain the purpose of a given tool, demonstrate its capabilities, and even convey its method of use [[112], [115], [168]]. There are essentially two kinds of animated icons: *icons that incorporate animated graphics* and *kineticons* [168]. This distinction is based on the fact that icons are generally composed of a bounding box which contains a pictogram. If the pictogram is animated, then it is an icon that incorporates animated graphics. If the bounding box is animated, then it is a kineticon.

Baecker *et al.* propose some high-level considerations for the design of the first kind of animated icons [112]; they identify ten basic ways in which they can be useful, and illustrate these with relevant questions.

Harrison *et al.* describe "six popular sources" of inspiration for the design of kineticons [168]. Five are references to real-world motions: *biological motion, gestures, organic motion, mechanical motion,* and *physics and natural effects*; which effectively rely on our ability to perceive causality and animacy in movement. The sixth is *cartoon conventions,* which are commonly accepted caricatures or exaggerations of real-world motions.

Overall, these general HCI design guidelines should certainly inspire the design of perceived affordances for online visualizations. However, general user interfaces are much more mature and familiar to the general public than visualizations, so the history and evolution of their perceived affordances should certainly be taken into account as well.

5.3 Suggested Interactivity

In light of the work presented in Section 5.2, we propose the following definition for Suggested Interactivity: *Suggested Interactivity is a set of methods for indicating that a graphical area can be interacted with by subtly directing a user's attention so as not to impede too heavily on this person's focus or on the rest of the interface design.* SI *cues* are then specific graphical elements or attributes that are used for suggesting interactivity.

While the concepts of perceived affordance and suggested interactivity are similar, we make the distinction based on the fact that the perception of affordance is generally related to design attributes of a unique/distinct interactive graphical object (*e.g.,* a widget), whereas SI is related to visual cues that do not necessarily pertain to an individual object: SI cues can be icons or text labels (*i.e.,* external objects) placed on top or next to an interactive area (*e.g.,* a visualization), which can be composed of several graphical objects. For example, it seems more appropriate to suggest the interactivity of a visualization as an *interactive area* (composed of visual marks, axes, *etc.*, which may all be interactive), rather than trying to design perceived affordances for each individual object it is composed of.

With these definitions in mind, we conducted a survey of a variety of highly interactive websites to identify how designers create and make use of SI cues for abstract interactive graphical objects and areas. We also surveyed several standard widgets, since, as we have mentioned above, new graphic standards tend to move away from the traditional embossments and drop shadows design. From this survey, we extracted a set of important dimensions for the design of SI cues, and constructed the design space presented in FIGURE 5.6. Note that while we focus on SI cues for visualizations here, this design space can be used to describe and generate SI cues for any kind of 'abstract' user-interface.

5.3.1 Survey and Design Space

5.3.1.1 Procedure and coding

We surveyed 230 HTML5 websites listed in [[5], [44], [75]], 150 data-journalism visualization websites listed in [46], and 2 of the *Gapminder* visualizations available for download [87]. This resulted in a total survey of 382 websites. We recorded all the different techniques used to suggest the interactivity of the webpages, or of specific graphical objects and areas within these (other than standard textual hyperlinks). Some websites did not include any SI cues, and many included similar ones. Overall, we identified 45 distinct cues, from which we extracted the following five main dimensions:

* **attractor:** the object that attracts attention to the interactive area;
* **animation:** the state of the attractor over time—note that in some cases the attractor can remain static;
* **trigger:** the event that initiates the animation;
* **visual attributes:** the specific visual variable(s) and/or mark(s) the animation is applied to; and
* **persistence:** the ongoing display or not of the cue once the interaction has been performed.

The **attractor** can either be the *object of interest*[1], *i.e.,* the interactive graphical object or area itself; or an *external object* (*e.g.,* an overlaid icon or text label). Its **animation**, when it exists, can be *staged, i.e.,* a predefined on/off *blink* or *interpolation*—which is either *unique* (one-shot) or *looped*; or *dynamic, i.e.,* dependent on specific 'page-level' user input (*e.g.,* mouse-move or mousewheel). The animation can be **triggered** by a *system-event*

1 Note that if the attractor is the object of interest and that this object is a unique graphical object, the Si cue can be considered a perceivable affordance.

(*e.g.,* pageload), or by a *user-event* (*e.g.,* mousemove, mouseover, click, or mousewheel), and can be applied to various **visual attributes** of the attractor, *i.e.,* to visual variables and/or to extra visual marks (similarly to hyperlinks, which use both). The **persistence** of the cue then determines whether it remains displayed after the intended interaction has been performed—in some cases it is removed immediately afterwards, as it can be considered that the user has discovered the interaction and will remember it throughout the rest of the session. Note however, that persistent SI cues can also be temporarily hidden, while the user is interacting with the interactive graphical object or area, or while this object is in focus. For example, the "play" button displayed on top of a video temporarily disappears while the user is watching the content.

To illustrate these dimensions, consider a standard hyperlink (even though we did not record these in our survey). The attractor is the object of interest, *i.e.,* the clickable text, to which no animation is applied. A visual variable and an extra visual mark are used, *i.e.,* color and underline. Finally, the cue is persistent, as it remains visible after a user clicks on the link, and later comes back to the webpage.

In addition to these main dimensions, we also coded the **intended interaction**, *i.e.,* the type of interaction the user is invited to perform; as well as the presence or not of **feedforward**, *i.e.,* a hint to the outcome of the interaction. We had also originally coded whether the SI cue was integrated to the visual narrative of the page, as is nicely done in [84] where a list of buttons is attached to a balloon that 'floats' against the background. However, this dimension turned out to be too complex, so we removed it. Finally, although it is not directly accounted for in our design space, we identified the distinction between icons that incorporate animated graphics and kineticons described in Section 5.2.4.3. In some cases, these were even combined. An attractor (object of interest or not) to which a staged or a dynamic animation is applied on an extra visual mark is generally an icon that incorporates animated graphics; and an external object attractor to which a staged or a dynamic animation is applied on a visual variable is generally a kineticon.

Our final design space is presented in FIGURE 5.6. Due to space limitations, we only reference one website per entry, but we provide a count (in brackets) of the number of websites that use each specific SI cue. In addition, as the attractor levels are mutually exclusive, we only included one row for this dimension: when an entry is coded with a black cell, the attractor is the object of interest, otherwise it is an external object. Similarly, the persistence and feedforward dimensions are binary, so a black cell indicates 'true.' We stress that the visual attributes we coded are only the ones to which the animation is applied. For example, an attractor may have a textual component, but if this component is not directly subject to the animation, it is not accounted for in our design space. Finally, we did not include *mousemove* in the intended interaction dimension, as we did not find, and could not think of any graphical object or area that simply relies on a mousemove to be interacted with.

In the following subsection, we present several specific cases we encountered in our survey, and discuss how these were fitted to our design space (when applicable).

5.3.1.2 *Discussion*

A majority of the SI cues we found (27/45) are applied to the object of interest; and in most cases (33/45), the type of animation is determined by what triggers it: staged animations are triggered by system-events (26/38) and dynamic animations are triggered by user-events (7/7). However, staged animations (12/38) can also be triggered by user-events. This occurs when SI cues are subject to sequenced interactions, *i.e.*, predefined linear series of interactions the user is invited to perform. Each interaction triggers the display of a new SI cue for the subsequent interaction. We highlighted these cases using red and blue cells. Sequences can focus on different interactive graphical objects or areas (blue cells) or on a same graphical object (red cells).

Sequenced interactions with different graphical objects or areas—To

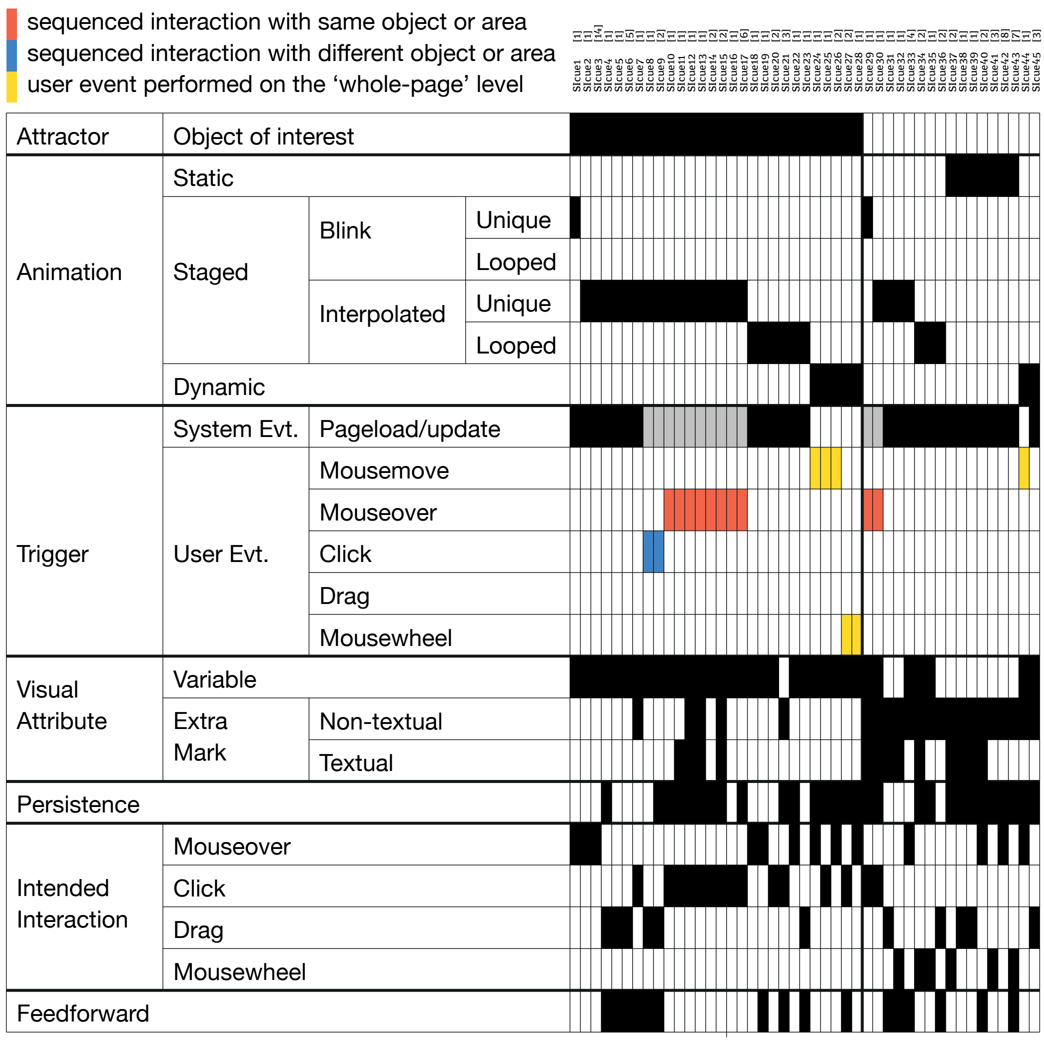


FIGURE 5.6: A design space of SI, based on our survey of 382 HTML5 and visualization websites. Due to space limitations, the table has been transposed, so entries are columns and dimensions are rows. The count of occurrences of each SI cue is shown in brackets, and the cue names refer to minified URLs (*e.g.,* SIcue1 can be retried at http://tiny.cc/SIcue1). Note that several URLs direct to the same websites, as these include multiple distinct SI cues. Finally, an interactive version of the design space with animated GIFs of each cue is available at [9].

illustrate this, suppose a user is required to click on a first graphical object GO1 before clicking on a second graphical object GO2. On page-load (a system-event trigger), a staged animation SI cue is applied to GO1, but no cue is applied to GO2. When the user identifies the SI cue and clicks on GO1 (*i.e.,* a user-event trigger), then a new staged animation SI cue is applied to GO2. An example of this is SIcue9: the user has to click on a 'play' button (GO1) to reveal an SI cue applied to a slider (GO2). We found this case hard to code as a user-event triggered SI cue, since the initial required interaction (performed on GO1) is unrelated to the second interactive graphical object (GO2). Thus, for simplicity we consider that any user-event that is not performed on the 'whole page' level (*e.g.,* mousemove or mousewheel—yellow cells) and that reveals an SI cue for an interactive graphical object (GO2) other than the one the user is already interacting with (GO1) triggers a new system-event (page update), which in turn triggers the animation of the SI cue. This way, we consider all staged animations to be triggered by system-events, which is why we have projected those that do not occur on pageload onto the system-event dimension in gray. As such, only dynamic animations are truly triggered by a user-event (which can occur anywhere on the page—yellow cells, with the exception of SIcue45, which we discuss in the Mouse cursor cases below).

Missing step cases—Sequenced interactions can also occur with a same graphical object. With regard to Buxton's *three-state model of graphical input* [127], we had expected that the SI cues for such sequences would follow a specific order, *i.e.,* hover first, then click, then drag (*e.g.,* SIcue18 then SIcue16). Standard hyperlinks do this, as they have an initial appearance that suggests a first interaction is possible, *i.e.,* mouseover, and a second appearance when a user hovers them that suggests another interaction is possible, *i.e.,* click. However, in several cases (*e.g.,* SIcue4, SIcue6) the SI cues skip the initial steps, for example inviting users directly to perform a drag interaction. While in many cases these steps are implicit, we stress that for abstract interactive graphical objects and areas, each step should be carefully considered, especially if the SI cue is not very *expressive*, *i.e.,* if it does not explicitly indicate what interaction is expected. For example, if a cue does not convey the idea of dragging and is not sequenced when a user hovers or clicks on the object or area, s/he may not move on to the next step, and will not understand the purpose of the interactive graphical object. Similarly, in some cases of sequenced interactions with different graphical objects, we found no SI cue for the first object the user must interact with in order to trigger the SI cue for the second object (*e.g.,* SIcue8, SIcue9).

Mouse cursor cases—In some cases, the mouse cursor itself is used as an SI cue (*e.g.,* SIcue45); the design of the cursor is modified to indicate a specific interaction the user can or should perform (*e.g.,* click-and-drag, with a hand icon and arrows indicating the direction in which the content can be dragged). These cases were particularly hard to code, as they can either be considered as inanimated (*i.e.,* no animation) attractors, or dynamic animation attractors. Here, we decided to code them as the latter, since obviously the mouse cursor is affected by the user-event mousemove, but it shows an exception since the cue is visible (or triggered) on pageload (a system-event). This kind of cue is often used when a 'whole page' interaction is dependent on clicking and dragging, like a swipe on a touch device.

Zoom widgets—Another ambiguous case was found in interactive maps with *zoom-inand-out* widgets (*e.g.,* SIcue43). These can either be considered as standard widgets that can be directly manipulated (by pressing on a + or - button, or by using a slider), in which case they are not really SI cues; or as cues for suggesting a mousewheel interaction, since zooming commonly relies on using the mousewheel. In our design space, we consider these widgets as the latter, and have coded them as external object attractors that invite users to scroll.

Combinations—Sometimes, both an object of interest and an external object attractor are used simultaneously. The intended interaction is often a mouseover, which highlights a specific region of the display and shows a tooltip (*e.g.,* [45]). Such combinations are very effective for providing feedforward. However, for the purpose of our design space, we have separated out the individual SI cues used in these combinations (*e.g.,* SIcue19).

Misleading cases—We encountered two cases in which SI cues were misleading. The first was in an NY Times graphic (SIcue5) that presents an interactive 3D map. The SI cue is a staged animation triggered on pageload that scales the map (the object of interest) into view, suggesting the possibility to zoom in and out. However, no mousewheel interaction is implemented for this purpose; only a click-and-drag is possible, but it only rotates the map in 3D space. The second was in another NYTimes graphic (SIcue7) that presents what seems to be a slider. The SI cue is a staged animation triggered on pageload that moves the slider thumb to a specific location, suggesting the possibility to click-and-drag it. However, the interactive area is actually a series of buttons which only allow for clicking; this SI cue provides a false idea of a continuous scale.

Edge cases—Finally, we encountered two edge cases, which required careful consideration for integration in our design space. The first was buttons, which like the zoom widgets discussed above, can either be considered as standard widgets, or as external object attractors. We consider that when the intended interaction can be performed in a region beneath or around a button (like in the case of a "play" button on top of a video where a user can click anywhere on the video frame to play the content), then the button is an external object attractor (*e.g.,* SIcue17). However, if the button requires clicking on directly to trigger something, then it is a widget, and we do not included it in our design space. The second edge case was found in [91], and is difficult to classify as an SI cue, as it relates to the layout of the interactive elements on the page. The interactive areas on this webpage are quite large and centered horizontally below the page-fold so that users, while scrolling with the mousewheel, accidentally end up hovering them; this then triggers another SI cue inviting users to click (SIcue11), creating a sequenced interaction with the same feature.

Interestingly, this example can be related to our hypothesis discussed at the end of Experiment 1 (see Section 5.1.1.3), where we considered that the layout of graphical objects and areas could be used to help discover interactions. However, we consider this **layout** dimension outside of our present scope, as it only relates to very specific edge cases.

5.3.2 Design Considerations for Visualizations

From this design space, we derived several considerations to operationalize the creation of SI cues for visualizations. While such low-level deconstructions are useful for describing existing designs, we find they are often too complex when it comes to creating new ones. As we are interested in suggesting the interactivity of charts embedded with text, we propose it is possible to use either the visualization itself as an attractor, *i.e.*, the object of interest, or an overlaid icon, *i.e.*, an external object. In this subsection, we discuss our considerations, and introduce several metaphors which we believe may assist designers.

5.3.2.1 Visualizations as attractors

Setting the visualization as the attractor limits the possible visual attributes and animations that can be used. Indeed, visualizations already depend on visual marks and variables to encode data, so those used for the SI cue should not overlap or interfere. In the case of an inanimated visualization (*i.e.*, an object of interest attractor with no animation), simply playing with free visual attributes [183] for the SI cue should be avoided as this could be considered more a stylistic choice than an invitation to interact. For example, using a red hue instead of a blue hue for a static animation bar chart is unlikely to be more effective for suggesting interactivity. Thus, using the visualization as the attractor requires applying a staged or a dynamic animation.

For staged animations, we propose the metaphor of *organic motion*, which consists in small repetitive animations, simulating the motion of organic processes to which we are "intimately familiar" [168]. Organic motions range from a heart beat to a timelapse of a blossoming flower. For dynamic animations, we propose the metaphor of *attractive motion,* which can consist in orienting, squeezing, or stretching a visualization, depending on how far the mouse cursor is from it. A typical example is SIcue24. Attractive motions range from a cat's head following the trajectory of a moth, to the orientation of a sunflower according to the sun, or of a metallic object attracted to a magnet. These different motions can then be applied to any visual attribute of the visualization, and can be persistent or not—although we do recommend stopping the animation when the user is interacting with the visualization, as this can be distracting and annoying.

5.3.2.2 *Icons as attractors*

External object attractors are generally icons, which may or may not be accompanied by text. These can be animated or not. Inspired by Baecker *et al.*'s considerations [112], we identify three kinds of icons: focal icons, identifier icons, and demonstrator icons. The first two generally use no animation, while the third uses staged animations. A *focal icon* is an icon displayed on top of a multimedia artifact like a video. When the artifact is *out of focus, i.e.,* when the user is not interacting with it, the icon is shown. When the artifact is *in focus,* the icon is removed. A typical example is the "play" button displayed on top of a video (*e.g.,* SIcue42). An *identifier* icon is usually an icon displayed next to an interactive area with a textual label indicating what should be done (*e.g.,* "navigate years" in [85]—SIcue40). However, identifier icons can also be used to replace the mouse cursor, in which case they are dynamically animated (as discussed in the Mouse cursor cases in Section 5.3.1.2). Finally, *demonstrator* icons are generally icons that incorporate animated graphics or kineticons (see Section 5.2.4.3); they show the user what to do in tutorial-like fashion (*e.g.,* SIcue31).

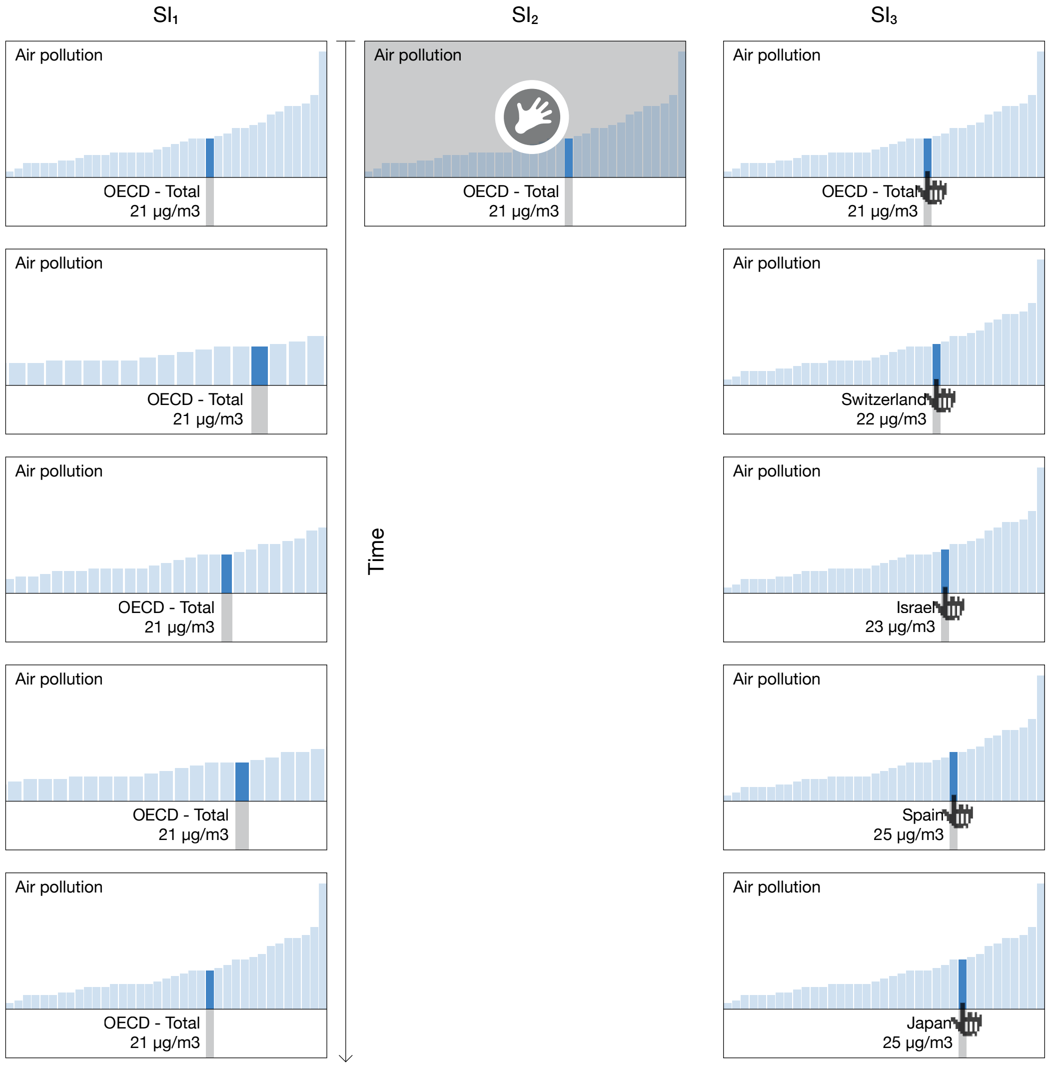


FIGURE 5.7: Three SI cues.

5.4 Testing Three SI Cues Applied to Bar Charts

To make an initial assessment of the effectiveness of SI cues applied to bar charts embedded with text, we generated a series of examples (available at [57]) and we tested the three we believe to be most representative of the diversity of our design space in a follow-up of Experiment 1 (see FIGURE 5.7). The first cue (SI1) we tested uses the object of interest (i.e., the visualization) as the attractor; the second (SI2) uses an external object attractor; and the third (SI3) uses a combination of both to provide feedforward (as is done in [45]—see paragraph on Combinations). In this section, we first describe the design of these cues and present their evaluation. We then provide some initial recommendations for design. Note that while it should prove interesting to test the full spectrum of variations that can be generated from our design space to see which are most effective, here we simply intend to assess whether SI cues actually have an effect on users' propensity to interact with charts.

5.4.1 Three SI Cue Designs

SI1 uses the visualization as the attractor and applies an organic motion to it (see [4]). We simulated a heart beat that first slowly stretches out, then bounces back into its original state. This staged animation is looped, triggered on pageload, and applied to the width of the bar chart. The cue is not persistent, *i.e.,*it is removed as soon as the chart is hovered over.

SI2 uses a focal icon as the attractor (see [48]). This respects the considerations mentioned in Section 5.3.2.2, and shows an open hand to suggest manipulation. The cue is persistent, as it is displayed again when the visualization is 'out of focus.'

Finally, SI3 uses both the visualization and a demonstrator icon as attractors (see [49]). For the visualization, we sequentially highlighted different bars using a looped blink animation. For the demonstrator icon, we mimicked a black pointer cursor to which we applied a looped staged animation to its horizontal position, simulating a brushing interaction. We also added an extra textual mark, *i.e.,*the label for the highlighted bar, to provide feedforward. This way, users have a sense of what they will find when performing the interaction. The cue is not persistent.

5.4.2 Experiment 4

To assess the effectiveness of these SI cues, we conducted a follow-up study on AMT. We reproduced Experiment 1 three times, respectively applying SI1, SI2, and SI3 to the bar charts. We then used the results of Experiment 1 as a baseline for comparison.

Three groups were tested, each with one of the SI cues, in a between subjects design. Group 1 (G1) was assigned SI1, Group 2 (G2) was assigned SI2, and Group 3 (G3) was assigned SI3. The scenario, tasks, and procedure were kept exactly the same as in Experiment 1.

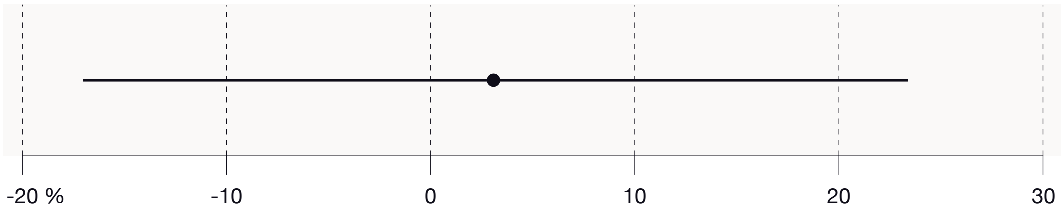
Hypothesis—We had the same simple hypothesis for each group:

* **H4:** participants will perform more brush interactions and decisive brushes when an SI cue is applied to the charts.

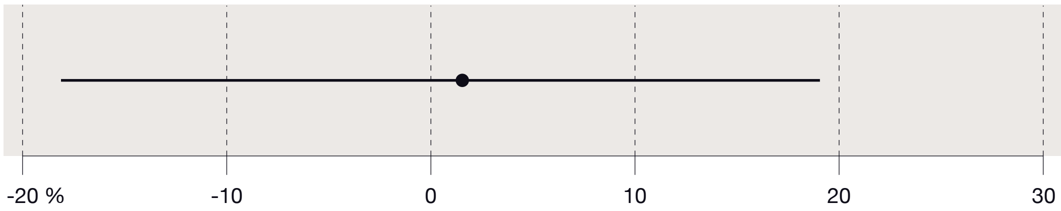
Participants—For each group, we recruited 40 different participants, making sure they had not participated in our initial studies. We retained the work of 33 participants in G1 whose total score was higher than 0; of 35 in G2; and of 40 in G3 (this was the only group in which all participants got scores above 0). All participants were native English speakers. We then coded the data in the same way as in our initial experiments.

5.4.2.1 Results

For each group, we calculated the difference between proportions of par-



(A) Difference between proportions of participants who performed brushing interactions in Experiments 1 and 4 G1 with 95% CI.



(B) Difference between proportions of participants who performed decisive brushes in Experiments 1 and 4 G1 with 95% CI.

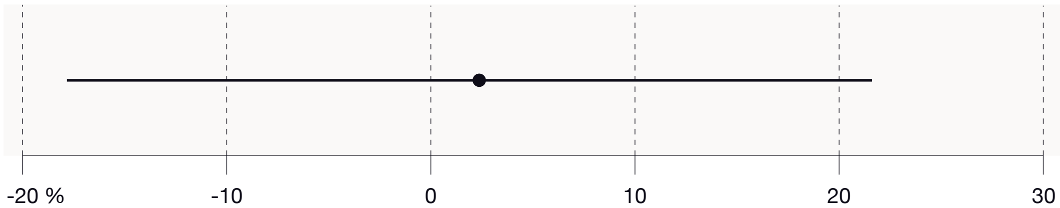
FIGURE 5.8: Results for G1.

ticipants who performed brush interactions in this experiment and in Experiment 1; we did the same for decisive brushes. Results are shown in FIGURE 5.8, FIGURE 5.9, and FIGURE 5.10 (for each group, respectively).

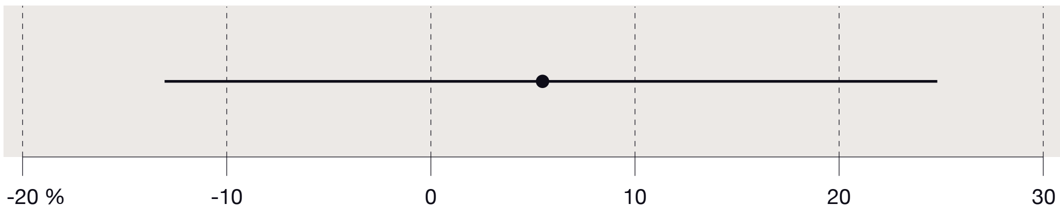
5.4.2.2 *Discussion*

Our results do not support **H4** for G1 and G2. The fact that the 95% CI are well below 0 shows no evidence that adding SI1 and SI2 to the charts enticed more participants to interact, or for that matter to use the charts for finding the answers (FIGURE 5.8 and FIGURE 5.9). However, **H4** is confirmed for G3: SI3 successfully incited more participants to interact, and to perform decisive brushes (FIGURE 5.10). We hypothesize that the success of SI3 is due to the fact that it provided feedforward. As shown in Experiment 2, people may need a short amount of time to learn how to use the charts, and we believe the feedforward helped them identify the benefit of interacting with the charts. However, it should be noted that this was

FIGURE 5.9: Results for G2.



(A) Difference between proportions of participants who performed brushing interactions in Experiments 1 and 4 G2 with 95% CI.

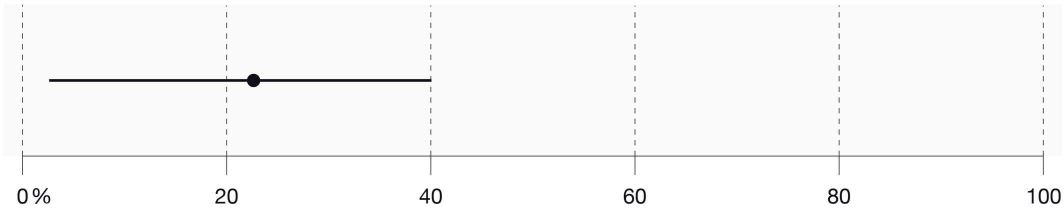


(B) Difference between proportions of participants who performed decisive brushes in Experiments 1 and 4 G2 with 95% CI.

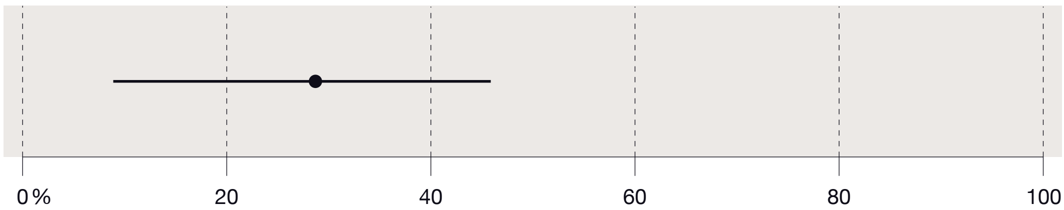
'heavy' design, as it required a combination two SI cues: one applied to the object of interest, and one applied to an external object attractor.

5.4.3 Initial Recommendations for Design

Overall, it seems that providing SI is necessary, especially when visualizations are designed for online audiences who may not be accustomed to the interactivity or different interaction techniques information visualizations may provide. While we have focused here on visualizations embedded with text, we strongly believe the same applies to other 'independent' online visualization applications. More people may expect these to be interactive, but there are no real conventions that can help them identify what can be done, *i.e., how* to interact with the display; and interactions with visualizations are generally more advanced than those required for other web-based media. Note that in our experiments, participants were only expected to discover a hover interaction, which can be considered as the



(A) Difference between proportions of participants who performed brushing interactions in Experiments 1 and 4 G3 with 95% CI.



(B) Difference between proportions of participants who performed decisive brushes in Experiments 1 and 4 G3 with 95% CI.

FIGURE 5.10: Results for G3.

simplest kind of interactions—as it can be performed 'involuntarily'—and even this was problematic.

Concerning design, while our results are only preliminary, we have found that our more subtle cues (*e.g.*, SI1 and SI2) were unhelpful, so similarly to hyperlinks, we believe a somewhat 'heavy' approach is necessary (*e.g.*, SI3). External object attractor SI cues can be combined with object of interest SI cues to provide feedforward, which can show a user what s/he can or should do with a visualization. A simple way of implementing such cues could be to create situated animated GIFs on the webpage (as done in [219]). However, we stress that this is not an immuable guideline: more work is needed on the evaluation of SI cues applied to visualizations. This could reveal that other/more subtle techniques and/or adjustments may be just as effective. For example, our initial experiments have led us to consider that for simple hover interactions, the position of the visualization in the webpage might have an effect. This should be properly tested.

Finally, it is still unclear how much interaction can or should be suggested. As a broader guideline, we encourage designers to aim for simple yet effective interaction techniques (*e.g.,* to implement complex infovis interaction techniques using only the array of standard interactions people usually perform on the web), as the current levels of *interaction literacy* or simply propensity seem generally low.

5.5 Conclusion

This chapter has focused on the **perceived interactivity cost**, and has addressed the following research question and design problem:

Q3: Do online users have a natural propensity to interact with visualizations—especially when these are embedded with text—and if not, how can we help these people detect the interactive potential of information visualizations?

I have shown that a majority of people lack initial propensity to interact with visualizations when these are embedded with text (**Q3**, part 1). To address this issue, Louis Eveillard, Françoise Detienne, Jean-Daniel Fekete, and I have introduced the concept of Suggested Interactivity, and have presented a survey of 382 HTML5 and visualization websites, in which we identified 45 distinct cues used to suggest the interactivity of abstract features like information visualizations. From this survey, we extracted a set of important dimensions for the design of SI cues, and have constructed a design space (**Q3**, part 2). We have then evaluated the benefit of using three representative SI cues for visualizations embedded with text, and have shown that an SI cue that provides feedforward can successfully entice more users to interact with charts, and thus can help them overcome the **perceived interactivity cost**. Our results also suggest that while certain people may lack initial visualization literacy, this problem can be rapidly overcome when questions and charts are highly-congruent.

We see four main avenues for future work on SI. The first concerns extending our evaluation of SI cues to establish whether other/more subtle cues can be as effective as SI3. The second concerns testing the cues we designed with different tasks. Those used in this article were very specific and focused, *i.e.,* fact-checking tasks. We intend to continue exploring the effectiveness of our SI cues for more open-ended tasks. The third concerns using SI cues for more complex visualizations, which are not embedded in text. While most visualizations still rely on widgets to perform dynamic queries, many now also propose direct manipulation techniques which are applied to the visual representation itself. These interactions need to be suggested to users, and we believe SI can be an effective means for doing so. Finally, the fourth concerns extending our design space by evaluating the *expressiveness* of cues, *i.e.*, how well they communicate the intended interaction to a user. This will require a more qualitative approach.

Ultimately, we consider the need for SI may simply be a transition phase: animated icons were necessary for a time in graphical user-interfaces, but have now mostly disappeared as users have become accustomed with such designs. Nevertheless, we believe it is important to guide users through this transition phase to accelerate the adoption of information visualization by *the* people.

Chapter 6

An Attempt to Provide Initial Incentives for Exploration: Using Storytelling to Engage Casual Audiences

In June 2014, Scott Murray released a podcast discussing an increasingly popular visualization format where authors tell a story first, before letting users dive into more details if they are interested (as described in Section 2.3.2.3) [105]. Such narrative visualizations are generally effective for conveying a message, or for persuading an audience. But can the use of *storytelling* in information visualization serve as means to get users interested in data? And can it be used to 'push' initial questions that may help them trigger an exploratory behavior? Although the work of visualization authors and journalists is important for creating the tools and context for information seeking, casual audiences should be provided with appropriate means and incentives for gaining insights and knowledge from data for and amongst themselves; and as discussed in the first two chapters of this thesis, this **initial incentive cost** could prevent people from engaging in data-explorations.

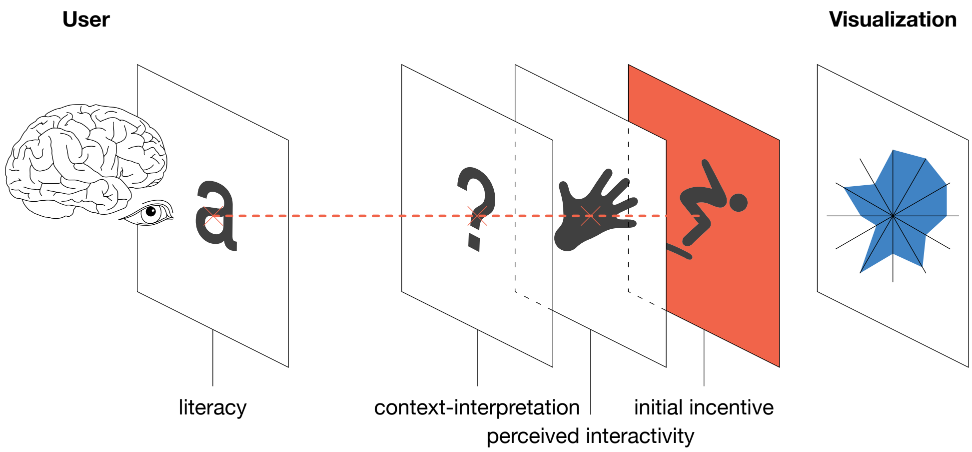


FIGURE 6.1: The **initial incentive cost**.

This chapter is based on a published paper entitled *Storytelling in Information Visualizations: Does it Engage Users to Explore Data?* [125], so any use of the term "we" refers to myself, Françoise Detienne, and Jean-Daniel Fekete. It focuses on the **initial incentive cost** by addressing the following research question:

Q4: Can providing initial incentives for exploration, *i.e.*, external motivations, in the design of visualizations trigger an exploratory behavior in casual audiences, and lead these people to engage in efficient personal data-explorations?

To answer this, we evaluate whether augmenting exploratory information visualizations with initial narrative visualization techniques and storytelling can help engage users in exploration. We use these techniques in an attempt to build interest in users by pointing out initial observations and insights from the data; and to provide initial incentives for exploration to users by 'pushing' unanswered questions and themes we had previously identified. We assess the impact of these techniques in three web-based field experiments, in which we compare user-behavior on a series of exploratory visualization webpages we designed that either included an initial story, or did not.

Many online data graphics use narrative design elements to explain a given dataset in a straightforward and compelling way. As discussed in Section 2.3.2.3, these explanatory graphics are preferable for data-journalism, as they have the advantage of exposing up-front what the main insights from the data are. However, most only provide limited interactivity [[238], Fig.7], which reduces the potential for personal extraction of insight. In essence and by definition, information visualization is interactive and exploratory. Thus, finding ways to make exploratory graphics more accessible and engaging to people is important, because, as mentioned in Section 1.1.1.3, if open data is to truly empower people, then these people should be able to use appropriate tools to gain their own insights and knowledge—not only that provided by journalists in articles written or designed from a specific perspective. Here, we explore the potential of narrative visualization techniques and storytelling to trigger this desired user-engagement. By *engagement*, we specifically mean a user's investment in the exploration of a visualization.

As such, the main contributions of this chapter are as follows:

* a first large-scale assessment of how users behave with different online information visualizations; and
* three assessments of the impact of 'pushing' observations, unanswered questions, and themes as initial incentives for exploration on user's behavior.

Note that we adopt an evaluation approach here, as our immediate motivation is to understand user-behavior with online information visualization, and to see whether augmenting these with popular narrative visualization techniques and storytelling can help engage users in exploration.

This chapter is organized in the following way. It begins with a background section that extends Section 1.2.3, Section 2.1.3, and Section 2.3.2.3; and presents previous behavioral metrics for measuring engagement. In Section 6.2, we then describe the design of our first experiment, and we present our analysis of user-behavior, and discuss our results. Section 6.3 presents the design of two follow-up experiments, for which we attempted to solve several problems identified in our first experiment, and discusses our results. Finally, in Section 6.4, we conclude with the implications of these results.

6.1 Background

As discussed in Section 1.2.3, Toms describes the process of information interaction [250] as a loop that cycles until a satisfactory amount of information is retrieved and integrated. According to Toms, users can initiate the interaction either by formulating a goal, or simply by deciding to examine a body of information. They then select or query a subset of this information, and scan it. When a cue is detected, they stop to examine the data, and if it is relevant, they extract and integrate it. Users can then recycle in multiple, nonlinear ways through each step.

However, while this model nicely conceptualizes the process of exploring an interactive-information-rich environment, it assumes that users have a relatively clear initial intent or questions in mind, and that they are capable of formulating appropriate queries using the interface. In the context of an online exploratory visualization, where viewers may not have specific background knowledge about the data or about visualization systems, we posit question articulation and data querying may be problematic. This is why designers and researchers [[77], [105], [179], [236], [238]] have suggested that storytelling can be used to trigger user-interaction and exploration, as it can provide the preliminary questions [238].

6.1.1 Using Narratives to Provide Initial Incentives for Exploration

Section 2.3.2.3 has presented several design considerations for narrative visualization. Segel & Heer suggest that an authoring segment can function as a “jumping off point for the reader’s interaction” [238].

However, while these frameworks are very useful for matters of design, it is still unclear whether the use of narrative visualization techniques in an introductory author-driven scenario can effectively lead to user engagement in a later more reader-driven scenario. Segel & Heer report some results of the deployment of a narrative visualization (*The*

Minnesota Employment Explorer) [238], but the intent of the study was to create and measure *social engagement* in the annotation of data with personal stories, rather than personal engagement in the exploration of provided data. Although we agree with Segel & Heer that an author-driven scenario is likely to help users articulate initial questions for exploration, we question whether it is sufficient for going "beyond those initial questions in depth and unexpectedness" [212].

6.1.2 User-Centered Metrics and Behavior

As discussed in Section 2.1.3, we adopt a behavioral and quantitative approach to measuring user-engagement here. While choosing appropriate metrics is essential for revealing underlying qualitative traits, these need to be related to a *goal,* and must be identifiable through different *signals* [233]. Here, our goal is to see whether augmenting exploratory information visualizations with initial narrative visualization techniques and storytelling can help engage users in exploration; we use low-level user-activity traces as signals, and we focus on analytic actions (which we refer to as *semantic operations*), and engagement—typically *depth of interaction,* which we interpret as the *number* of interactions a user performs that have a direct and perceivable impact on the display.

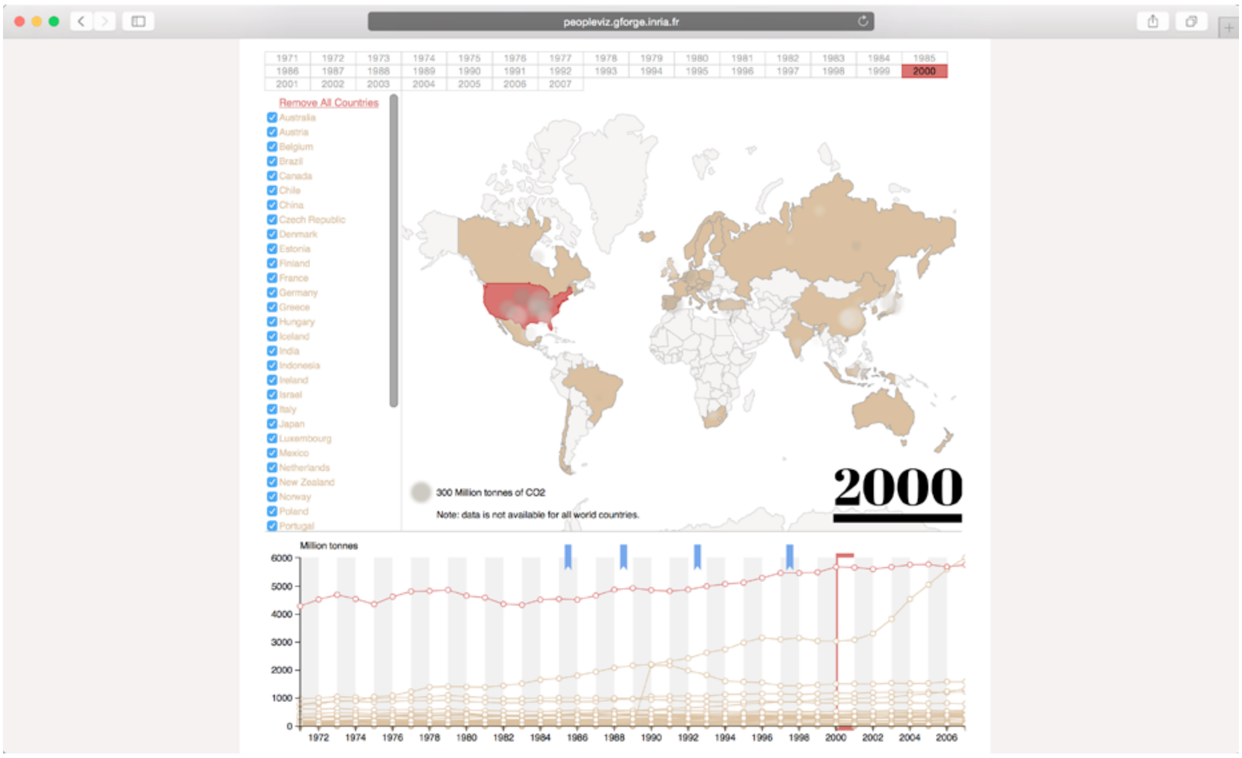


FIGURE 6.2: The *CO2 Pollution Explorer*—Explore section.

6.2 Case 1: The CO2 Pollution Explorer

In the rest of this chapter, we describe the design and results of our three field experiments. For each, we created a specific exploratory visualization webpage with two versions: one that included an introductory *narrative component*, which told a short *story* about the topic and context of the data, provided initial insights and unanswered questions, and introduced the different visual encodings; and another that did not. Each version, which we respectively refer to as the *Storytelling (ST)* version and the *no-ST* version, was alternately assigned to new browser connections; returning connections were reassigned the same version using a Cookie. Thus, our experimental design was between-subjects.

By comparing user-behavior between versions, we seek to determine whether augmenting such a visualization with an introductory story can help engage users in exploration. Our first field experiment was conducted with the *CO2 Pollution Explorer* [42], which was first published in English on visualizing.org, a popular online visualization gallery, then in French on Mediapart, a popular French news and opinion outlet. The visualization was referenced as one of the "Best of the visualisation web... January 2014" on visualisingdata.com, and was curated in the "Visualizing Highlights: March 2014" on visualizing.org. It was also picked up by bloggers on reddit.com, citylab.com, various other sites, and social media. Altogether, the webpage received roughly four thousand unique browser connections between January and June 2014.

6.2.1 Design

The *CO2 Pollution Explorer* (FIGURE 6.2) presented a dataset on the yearly evolution of CO2 emissions in different countries of the OECD. The two main graphical components were the *CO2 Pollution Map* (presented in Chapter 4), which showed the emission of each country as an animated smoke cloud, and a line graph, which showed the evolution of emissions over time. The narrative component in the ST version was designed as a heavily author-driven (although "user-directed") slideshow with messaging, that included five stimulating default views (or *sections*—see Appendix M). These were sequenced using a set of stepper-buttons, which triggered various animated transitions [[238], Fig.7]. Each section followed the general layout shown in FIGURE 6.3 (A), and interactions were limited to clicking on the stepper-buttons and hovering over the graphic—this displayed an inspector with country names and/or total CO2 emissions. This design was directly inspired by many well accepted and highly acclaimed NYTimes graphics [65]. After the narrative component, the webpage 'opened up' (similarly to the Martini Glass structure) to an *Explore section*, which included only a small amount of messaging, and introduced

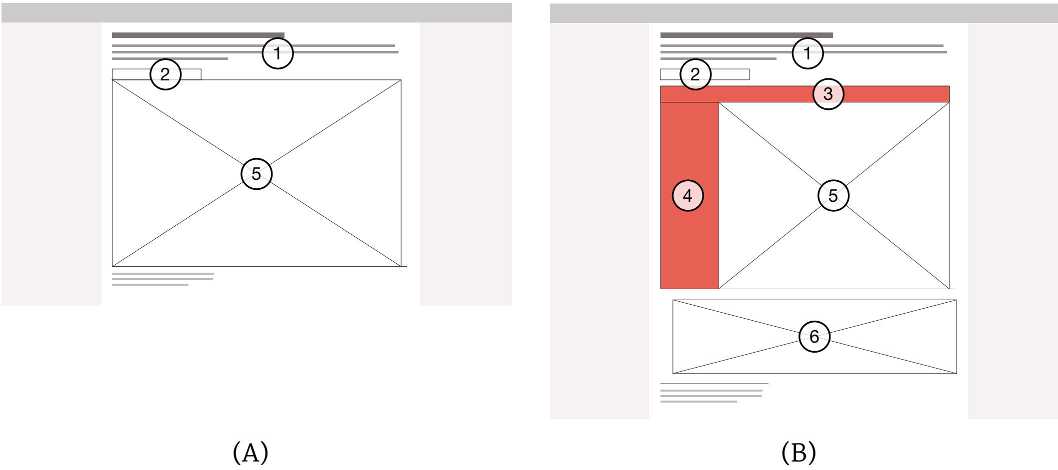


FIGURE 6.3: General layouts for sections in the narrative component (A), and for the Explore section (B). 1) short descriptive paragraph (messaging), 2) stepper-buttons (to navigate between sections—only in the ST version), 3) query-buttons, 4) list of country names and query checkboxes, 5) main graphic, and 6) secondary graphic. Additional interactive features in the Explore section are highlighted in red.

several extra interactive features that visitors could freely use to explore the dataset. This section followed the layout shown in FIGURE 6.3 (B), and was what visitors assigned to the no-ST version were shown.

6.2.1.1 *Metrics*

Using the categories of interaction described by Yi *et al.* [274] and inspired by Gotz & Wen's analytic actions [85], we created the following taxonomy of semantic operations users could perform with the *CO2 Pollution Explorer*. Each level corresponds to one or several low-level interactions with specific features of the interface (presented in brackets).

* **inspect:** show the specifics of the data [*hover line graph, hover line graph dot, hover map*];
* **connect:** show related items [*hover list*];

* **select:** mark something to keep track of it [*click line graph, click line graph dot, click map, click list label*];
* **filter:** show something conditionally [*click list checkbox, click "Show All Countries/Remove All Countries" button*];
* **explore:** show something else [*click query-button*]; and
* **narrate:** show a different section [*click stepper-button*].

To make sense of the four thousand sessions we collected, we performed some initial filtering and manipulations. 1) Since the webpage was designed for desktop browsers, we removed all mobile device connections. While displayable on such devices, the visualization offered certain interactive capabilities that touch displays do not handle (*e.g.,* hovering). In addition, mobile device displays are generally smaller than desktop ones, and we could not assert that the visualization fully fitted the screen resolution, or that if it did, it would not be too small to read and interact with. 2) While several sessions were those of returning browser connections, we considered each of them individually. *Return* is a good indicator for user-engagement [233]. However, analyzing aggregated sessions would have created major outliers for other metrics such as uptime or depth of interaction. In addition, while it may be conceivable that certain users opened the page, read or explored it for a moment, then turned to another activity, only to later come back and finish their exploration, our traces do not show whether users remembered what they had previously done or that they were not distracted by some external factor. In line with this, we also set a ten minute threshold for inactivity within sessions. Each session in which two consecutive traces were separated by ten or more minutes were split in two. 3) We removed all browser connections that had arrived to the webpage through social media and personal blogs. This allowed us to categorize two different visitor populations: on the one hand, we had visitors coming from visualization gallery websites, which we consider to be a *visualization-savvy* population, and on the other hand, we had visitors coming from Mediapart, which we consider to be an *information-savvy* population, but with *a priori* no particular interest in visualizations—since

Mediapart very rarely publishes interactive data-graphics. 4) Finally, several of the browser connections in our traces were ours, as we had originally used the live version for debugging and demoing. Unfortunately, we had no direct method for removing these sessions, since UUIDs were random and anonymous. However, we never communicated the URL directly, and it was hard to guess or remember. Therefore, to filter out our own sessions, we removed all connections that had no previous page URL. In the end, this procedure resulted in a subset of **2975 sessions.**

To obtain practical metrics, we coded the visitor-activity traces in the following way: 1) We attributed session IDs to each returning session and computed the uptime of all sessions. We also separated out the time visitors in the ST version spent in the narrative component and the time they spent in the Explore section. 2) We counted the total amount of click and hover interactions, and extracted all *meaningful interactions.* We define *meaningful hover interactions* as *hover interactions that affect the display (e.g., an inspector overlay) and that last longer than 250ms, so that the user can perceive its effect on the display;* and *meaningful click interactions* as *click interactions that occur on interactive features of the display (i.e., not random clicks anywhere on the display).* We then added these meaningful interactions to get a total meaningful interactions count per session. 3) We separated out the different semantic operations, and we repeated the interactions coding procedure for the Explore section alone (in the ST version). This provided us with comparable values for identical settings in both versions. 4) Finally, we coded the sections visitors inspected in the ST version in a dichotomous way: inspected sections were coded 1 and all others 0; and we controlled for linear sequencing of these sections by looking for the pattern [1, 2, 3, 4, 5, Explore] and coding 1 when matched, and 0 otherwise.

6.2.2 Hypotheses

Our analysis was driven by two qualitative hypotheses. The first was that **the narrative component should effectively immerse users in the ST version**, resulting in the fact that they should read through the whole story at least once in a linear fashion, and the second, that **the presence of this narrative component should effectively engage users in the exploration of the data**, resulting in higher user-activity levels in the Explore section of the ST version than in the whole no-ST version. However, verifying such qualitative hypotheses in a web-based field experiment is impractical. Therefore, we operationalized them with the following six quantitative hypotheses:

* **H1.1** (whole webpage): Visitors in the ST version spend more time on the webpage than those in the no-ST version;
* **H1.2** (whole webpage): Visitors in the ST version perform more meaningful interactions with the webpage than visitors in the no-ST version;
* **H2.1** (ST version only): A majority of visitors in the ST version inspect all six sections of the webpage;
* **H2.2** (ST version only): A majority of visitors in the ST version inspect the six sections in a linear fashion;
* **H3.1** (Explore section only): Visitors in the ST version spend more time in the Explore section than visitors in the no-ST version; and
* **H3.2** (Explore section only): Visitors in the ST version perform more semantic operations in the Explore section than visitors in the no-ST version.

We conducted separate analyses for the two populations mentioned above; each was composed of three phases. First, we looked at the general differences between the ST and the no-ST versions. Then, we inspected the ways in which visitors in the ST version inspected the narrative component. Finally, we compared the ways in which visitors behaved in the Explore section between versions.

6.2.3 Results

In the following subsections, we present the results for the information-savvy population (**1270 sessions**). In the subsequent Discussion section, we simply report the similarities and discrepancies we found with the visualization-savvy population (**1705 sessions**).

As in Chapter 5, we base all our analyses and discussions on estimation, *i.e.*, effect sizes with confidence intervals (95% CI). Effect sizes are reported as ratios between values for the ST version and values for the no-ST version. All point estimates and 95% CI are based on 10, 000 percentile bootstrap replicates of the statistic applied to the data [130].

6.2.3.1 *Whole Web-Page Analysis*

The first part of our analysis focused on standard aggregated Web analytics (*i.e.*, total *uptime* and *click-count*). We began by inspecting the webpage's uptime in both versions. We applied a logarithmic (log) transformation to the data in an attempt to normalize their distributions. Nevertheless, the dashed histogram in FIGURE 6.4 shows a bimodal distribution, and the one in FIGURE 6.5 is skewed. To explain this, we looked at the day of the week and the time of the day at which visitors connected to the webpage, expecting that during working hours, sessions would be shorter. This was not the case. Pursuing, we considered that the abnormality of the distributions might be due to bouncing behaviors. The *Google Analytics Help page* [104] defines *bounce rate* as *the percentage of single-page visits.* While this definition is not directly applicable in our case, since we use a single dynamic page, we interpret this metric as *the percentage of visitors who perform no click interaction on the page*—since seeing different pages of a website boils down to clicking on a series of hyperlinks[1]. We emphasize that this interpretation strictly concerns the absence of click interactions, since hover interactions may be incidental.

1 In some cases, the bounce rate is not a 'negative' metric: visitors may just find the information they need on the first page without having to perform an interaction. However, in our case, the amount of information readily available on page-load is quite low.

[step 1]—19.2% of sessions in the ST version and **14.6%** in the no-ST version showed a bouncing behavior. The *geometric mean* (GM) durations of these sessions were **9.5 seconds** (s) and **17.9s**, respectively. We expected this to be the result of returning users who had already read the story and/or seen the visualization. However, the return IDs showed that **71.9%** (88/122) of bounces occurred in first-time connections in the ST version, and **70.4%** (65/93) in the no-ST version.

[step 2]—We removed all bounced sessions from further analysis, and plotted the uptime distributions again. The solid histograms in FIGURE 6.4 and FIGURE 6.5 show that they are now near-normal.

[step 3]—We then compared uptime in both versions. FIGURE 6.6 provides evidence that visitors in the ST version spent more time on the webpage (GM = **123.8s**, 95% CI [115.3, 132.9]) than visitors in the no-ST version (GM = **101.6s**, 95% CI [101.6, 117.1]), since the ratio is above 1.

[step 4]—Next, we turned to the number of meaningful interactions. Visitors performed on average **42.7 meaningful interactions**, 95% CI [39.3, 46.3] in the ST version, and **43.7**, 95% CI [40.3, 47.3] in the no-ST version (FIGURE 6.7). This provides no evidence of a difference between versions.

[step 5]—We then conducted separate comparisons of the meaningful hover and click interactions. FIGURE 6.7 provides little evidence that visitors in the no-ST version performed more meaningful hover interactions. However, it provides good evidence that visitors in the ST version performed more meaningful click interactions.

6.2.3.2 *Narrative Framework Analysis*

The second part of our analysis focused on the narrative framework, and the way visitors in the ST version navigated through the different sections

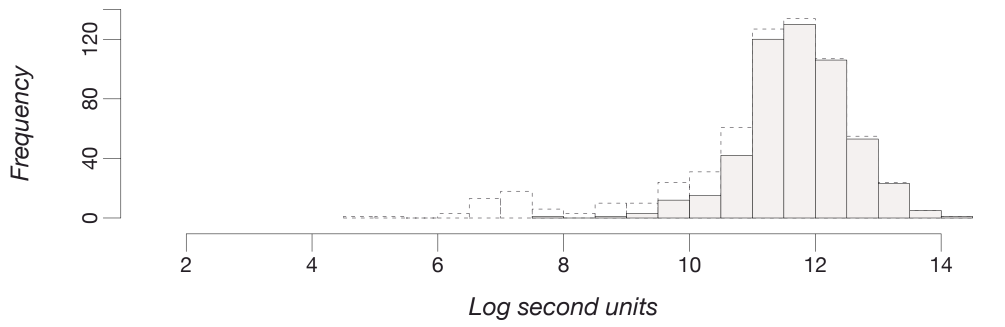


FIGURE 6.4: Log uptime distribution in the ST version.

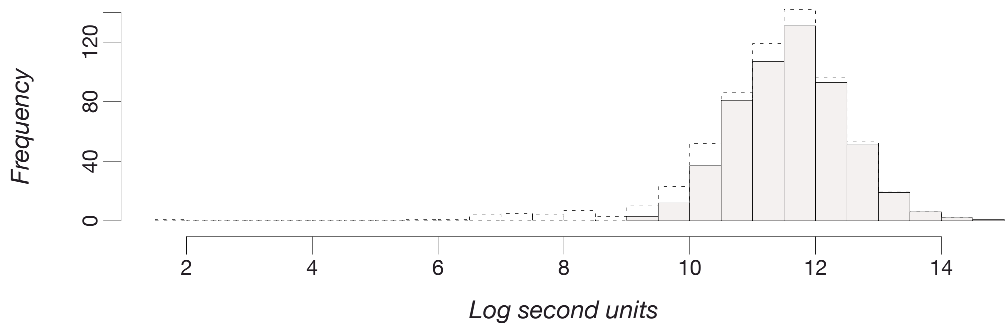


FIGURE 6.5: Log uptime distribution in the no-ST version. In both of these figures, dashed histograms represent distributions before removal of bounced session, and solid histograms represent distributions after removal.

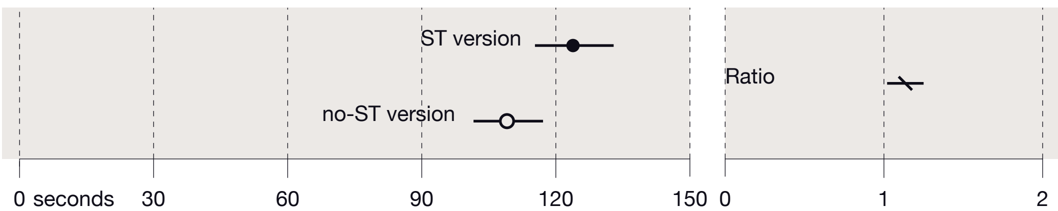


FIGURE 6.6: [step 3] Geometric mean uptime with 95% CI and ratio.

of the narrative component and the Explore section.

[step 6]—We began by looking at the number of sections visitors had inspected. In all sessions, visitors saw more than one section; in **77.5%**, they saw the Explore section; and in **71.7%**, they saw all six sections. Similarly to the bounce rate, we expected that the sessions in which visitors did not inspect all sections would be returning visits, where visitors would have already seen some (if not all) of the content. However, the return IDs showed that **86,2%** (125/145) of these sessions were first-timers.

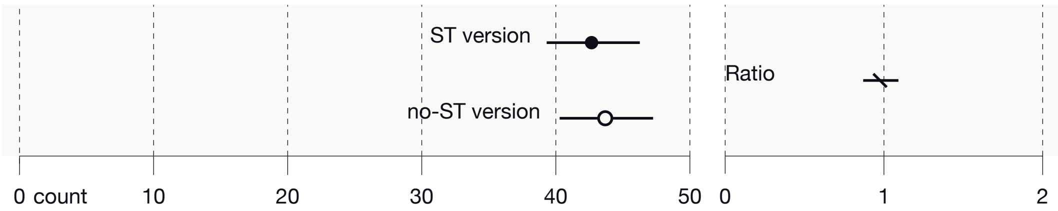
[step 7]—We removed all sessions in which all six sections had not been inspected from further analysis, and turned to the number of sessions in which the narrative component and the Explore section had been inspected in a linear fashion. Only **35.4%** (130/367) met this requirement.

6.2.3.3 *Explore Section Analysis*

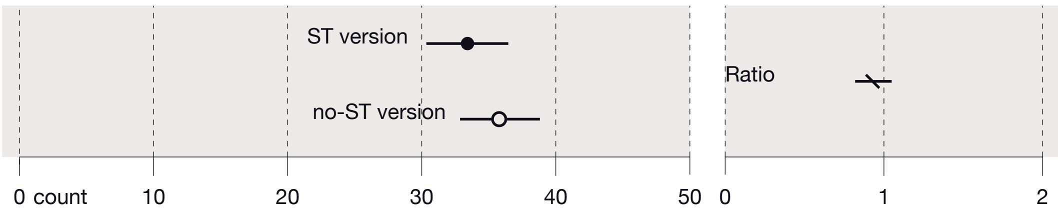
The last part of our analysis focused on comparing visitors' behavior in the Explore section between versions. Remember that in the no-ST version, visitors were only shown the Explore section, so the time they spent and the interactions they performed in this section are the same as those for the whole webpage [steps 1 through 5].

[step 8]—We began by looking at the time visitors spent in the Explore section. These durations were normally distributed (once log transformed) for both versions, and their geometric means and 95% CI (FIGURE 6.8) provide good evidence that visitors in the noST version spent twice as much time in Explore section as visitors in the ST version did (**108.8s > 54s**).

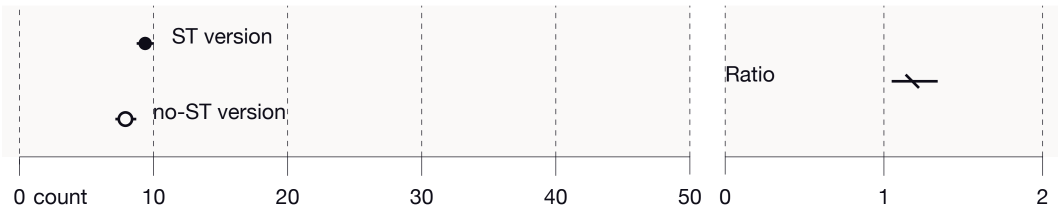
[step 9]—Next, we compared the amount of meaningful interactions. FIGURE 6.9 provides good evidence that visitors in the no-ST version performed more hover and click interactions than visitors in the ST version.



(A) Total number of meaningful interactions with 95% CI and ratio.



(B) Number of meaningful hover interactions with 95% CI and ratio.



(C) Number of meaningful click interactions with 95% CI and ratio.

FIGURE 6.7: Results for [steps 4 and 5].

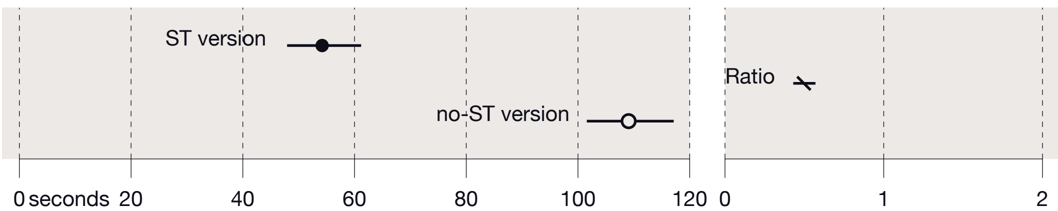
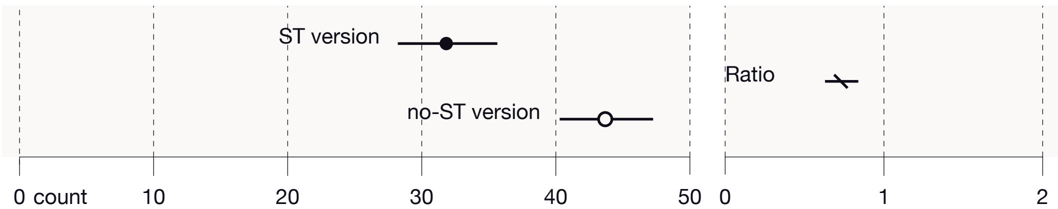
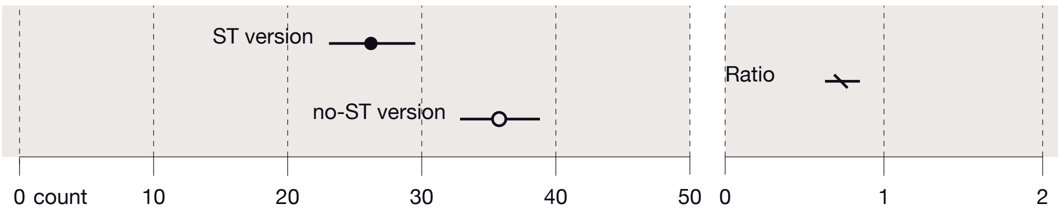


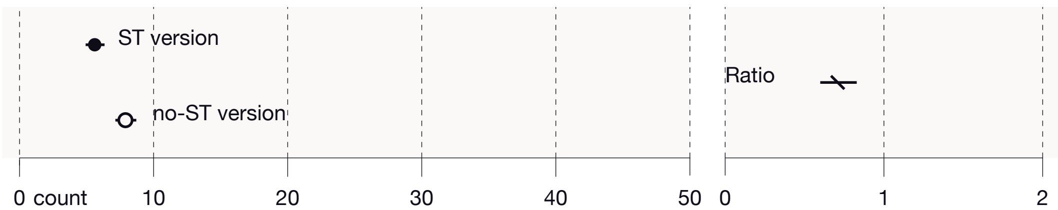
FIGURE 6.8: [step 8] Geometric mean time spent in the Explore section with 95% CI and ratio.



(A) Total number of meaningful interactions with 95% CI and ratio.

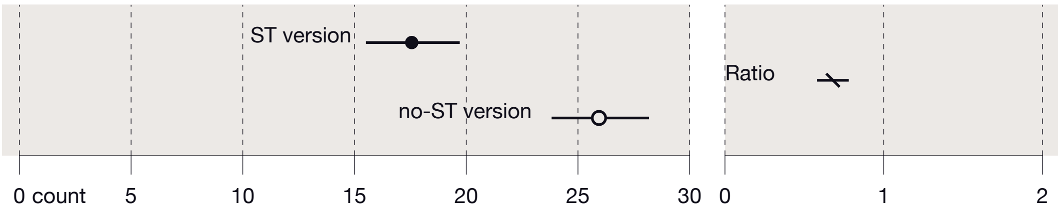


(B) Number of meaningful hover interactions with 95% CI and ratio.

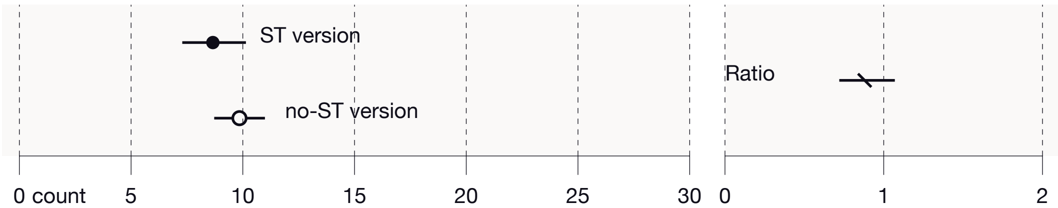


(C) Number of meaningful click interactions with 95% CI and ratio.

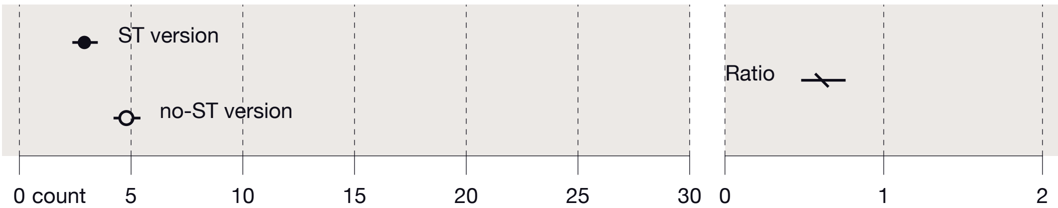
FIGURE 6.9: Results for [step 9] (in the Explore section alone).



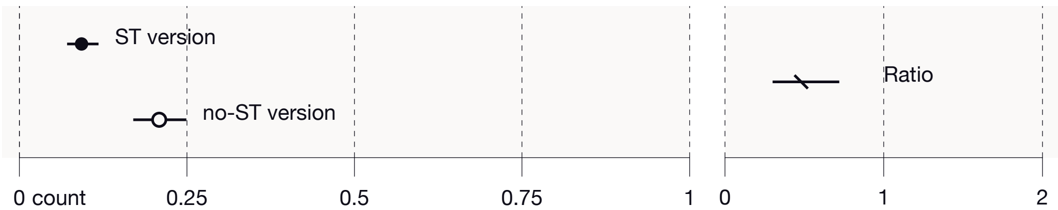
(A) Number of *inspect* operations with 95% CI and ratio.



(B) Number of *connect* operations with 95% CI and ratio.

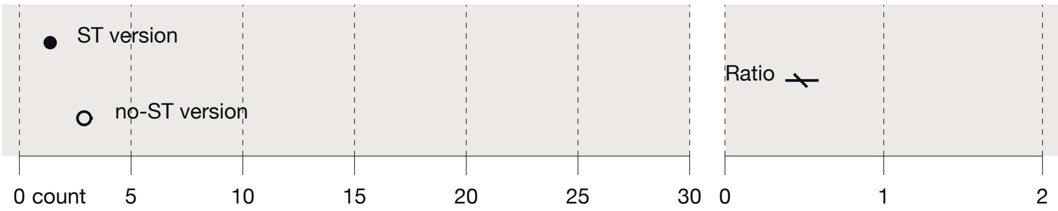


(C) Number of *select* operations with 95% CI and ratio.



(D) Number of *filter* operations with 95% CI and ratio.

FIGURE 6.10: Results for [step 10] (in the Explore section alone—Figure is continued on next page).



(E) Number of *explore* operations with 95% CI and ratio.

[step 10]—After that, we turned to the semantic operations visitors performed. We did not consider *narrate* operations here, as they were not available in the no-ST version. A summary is given in FIGURE 6.10. All 95% CI but the ones for *connect* operations provide good evidence that visitors in the no-ST version performed more semantic operations than visitors in the ST version. The figure also provides good evidence that in both versions, visitors mainly performed *inspect* operations. However, the most surprising finding here is that nearly no visitor at all performed *filter* operations.

6.2.4 Discussion

H1.1 is confirmed by our results [step 3]. However, the 20% bounce rate in the ST version [step 1] might indicate a certain miscomprehension of the purpose of the stepper-buttons: visitors may not have realized that it was possible to display other content. While we did not pilot-test the usability of these buttons *per se,* we did show the ST version to several people before publishing the webpage (including our editor at Mediapart), and the stepper was not an issue. Thus, another explanation, when considering the 15% bounce rate in the no-ST version and the fact that most bounces in both versions were first-time sessions, might simply be that visitors had trouble displaying the webpage; one visitor reported this, and attributed it to the browser extension *Ad Block Plus* (ABP).

H1.2 is only partially confirmed, as visitors in the ST version only performed more click interactions [step5]. While these two conclusions seem rather obvious (since there was more content in the ST version), and although it may be argued that the two versions are difficult to compare at this level, we posit these results can be valuable for *e.g.,* a publisher, who may simply want to know what format will increase the uptime and click-count of an article.

H2.1 is also confirmed [step 6]. To estimate whether these visitors actually read the textual content of the narrative component, we conducted a *post-hoc* analysis to determine their word per minute (wpm) score. *wpm* is a standard metric for reading speed [133], and according to [231], the average French reader's score is between 200 and 250 wpm. Visitors spent roughly 78s (GM) in the narrative component, where there were altogether 269 words to read. Their average wpm is thus **207**, which makes it plausible to assume that they read the story, even if they spent extra time inspecting the graphics.

H2.2 however, is not confirmed [step 7]. This reinforces our idea of a possible miscomprehension of the purpose of the stepper-buttons. *These may not have been explicit enough to convey the idea of a linear narrative* (P1)(2).

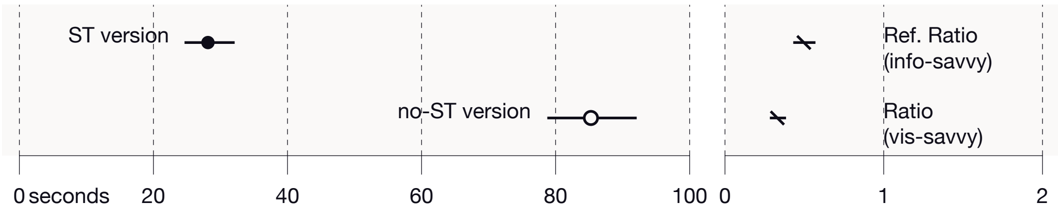
H3.1 and **H3.2** are not confirmed either [steps 8 and 10]. It should be noted however, that the interaction counts in the no-ST version are likely to include erroneous interactions, *i.e.,* interactions that visitors performed just to get used to the interface, without any specific analytical intent. Nevertheless, we consider these negligible, since the only operations that visitors in the ST version could have gotten 'used to' in the narrative component were *inspect* operations; and, even though the evidence is low, it seems visitors in the no-ST version performed altogether more hover interactions [step 5].

Overall, these results invalidate our qualitative hypotheses: the narrative component did not immerse visitors in the way we expected it to, since they did not inspect the story in a linear fashion; and it did not increase visitors' engagement in exploration in the Explore section. This

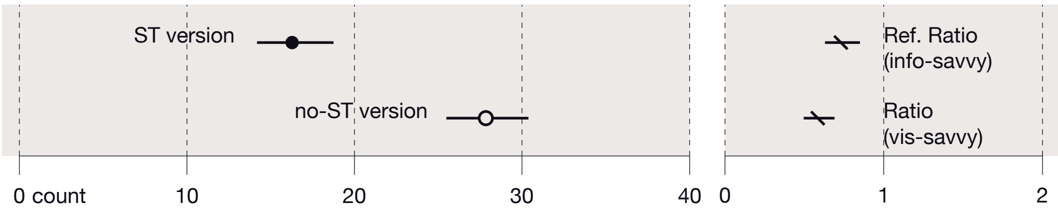
2 We point out possible design or usability problems uncovered by our analysis in this section, and discuss how we fixed them for our follow-up experiment in the next section.

suggest that augmenting an exploratory visualization with initial narrative visualization techniques and storytelling does not increase user-engagement in exploration. Nevertheless, it does not mean that visitors in the ST version retrieved less information from the webpage than visitors in the no-ST version did: our results simply do not account for this. In fact, since the narrative component provided several important insights, it is possible that visitors in the ST version actually got more information out of the webpage. However, this was information we provided, not personal insight. As visitors in both versions mainly performed *inspect* operations [step 12], it seems that the Scan pattern was predominant [166], and that visitors' main analytical intent was to simply compare the specific amounts of CO2 emitted by the countries displayed by default at a single point in time. A possible explanation for this limited exploratory behavior after having inspected the narrative component in the ST version is that *visitors may have considered the information presented in the story to be sufficient* (P2). Alternatively, it may be that *our design of the narrative component was not compelling enough to help them articulate questions about the data* (P3), and *did not sufficiently 'train' them to use the interactive features of the Explore section* (P4). It is also possible that *visitors did not perceive the dataset as being rich enough for them to spend extra time exploring it* (P5)—one visitor commented that "the graphic is interesting, but it lacks a key piece of information necessary to a political solution for the reduction of greenhouse gases: the emission rate per capita!" [[42], on Mediapart]. Visitors may have indeed had too much *a priori* knowledge of the topic. A final explanation we can think of is that the visualization itself may not have been appealing enough. Toms reports that "the interface must rationally and emotionally engage the user for satisfactory resolution of the goal. [...] content alone is not sufficient" [250]. *The interactive features of the CO2 Pollution Explorer may not have been explicit enough or may have been perceived as too limited* (P6)—as suggested by the general absence of *filter* operations [step 11].

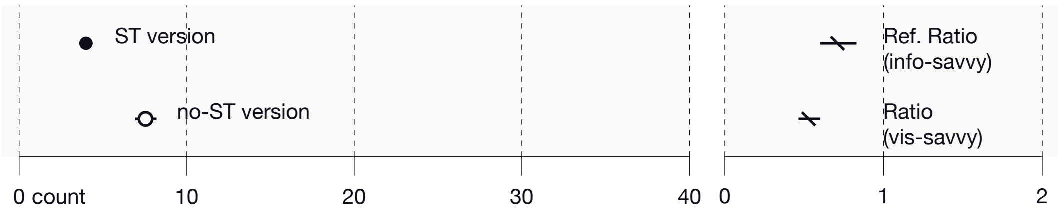
Nevertheless, the webpage did generate some interesting debate in the Comments sections of the websites it was picked up by—typically on citylab.com, visitors discussed "who's responsible for cleaning up our



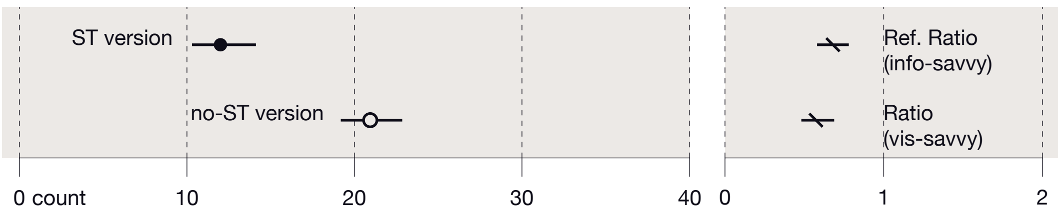
(A) Geometric mean time spent in the Explore section with 95% CI and ratio.



(B) Number of meaningful hover interactions with 95% CI and ratio.

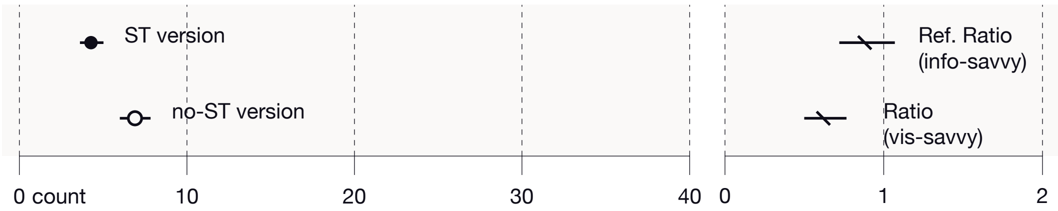


(C) Number of meaningful click interactions with 95% CI and ratio.

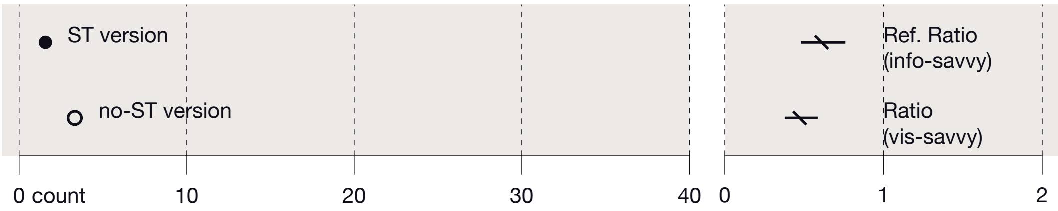


(D) Number of *inspect* operations with 95% CI and ratio.

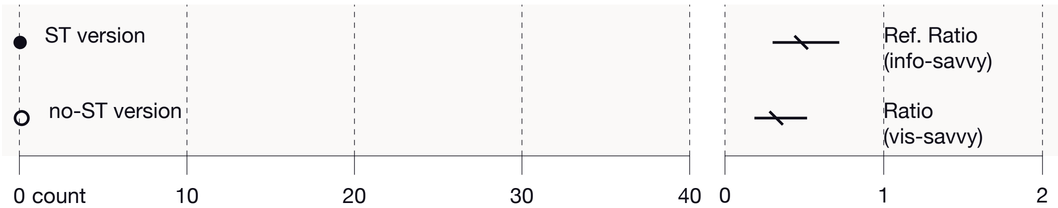
FIGURE 6.11: Activity comparisons between versions for the visualization-savvy (vis-savvy) population in the Explore section alone. Ratios for the information-savvy (info-savvy) population are also shown for reference (Figure is continued on next page).



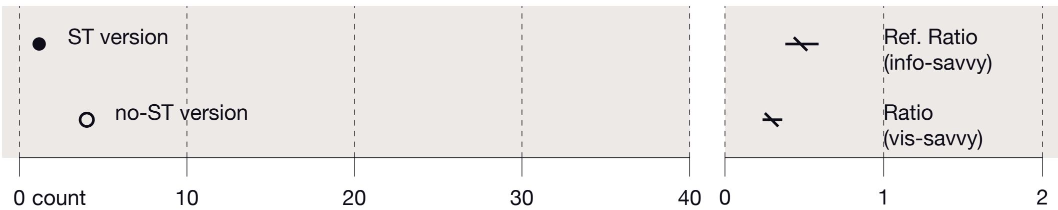
(E) Number of *connect* operations with 95% CI and ratio.



(F) Number of *select* operations with 95% CI and ratio.



(G) Number of *filter* operations with 95% CI and ratio.



(H) Number of *explore* operations with 95% CI and ratio.

past?", as well as possible solutions for the future, such as "a Manhattan Project for clean energy production" [[42], on citylab. com]; but unfortunately, it is impossible to tell which version these people had seen.

Finally, while we had expected that the behavior of the visualization-savvy population would be different, especially concerning interactive-behavior, it was overall very similar; uptime was slightly shorter and interactions count smaller, but the general trends were the same—as illustrated by the ratio comparisons in FIGURE 6.11, with the minor exception of the number of *connect* operations, for which there is good evidence here that visitors in the no-ST version performed more.

FIGURE 6.12: The *Economic Return on Education Explorer* (the richer visualization)—Explore section.

6.3 Cases 2 & 3: The Economic Return on Education Explorer and the Nuclear Power Grid

To ensure these unexpected results were not confounded by the possible design or usability problems pointed out in the previous section, we conducted a follow-up study with two other exploratory visualization webpages—the *Economic Return on Education Explorer* [107] and the *Nuclear Power*

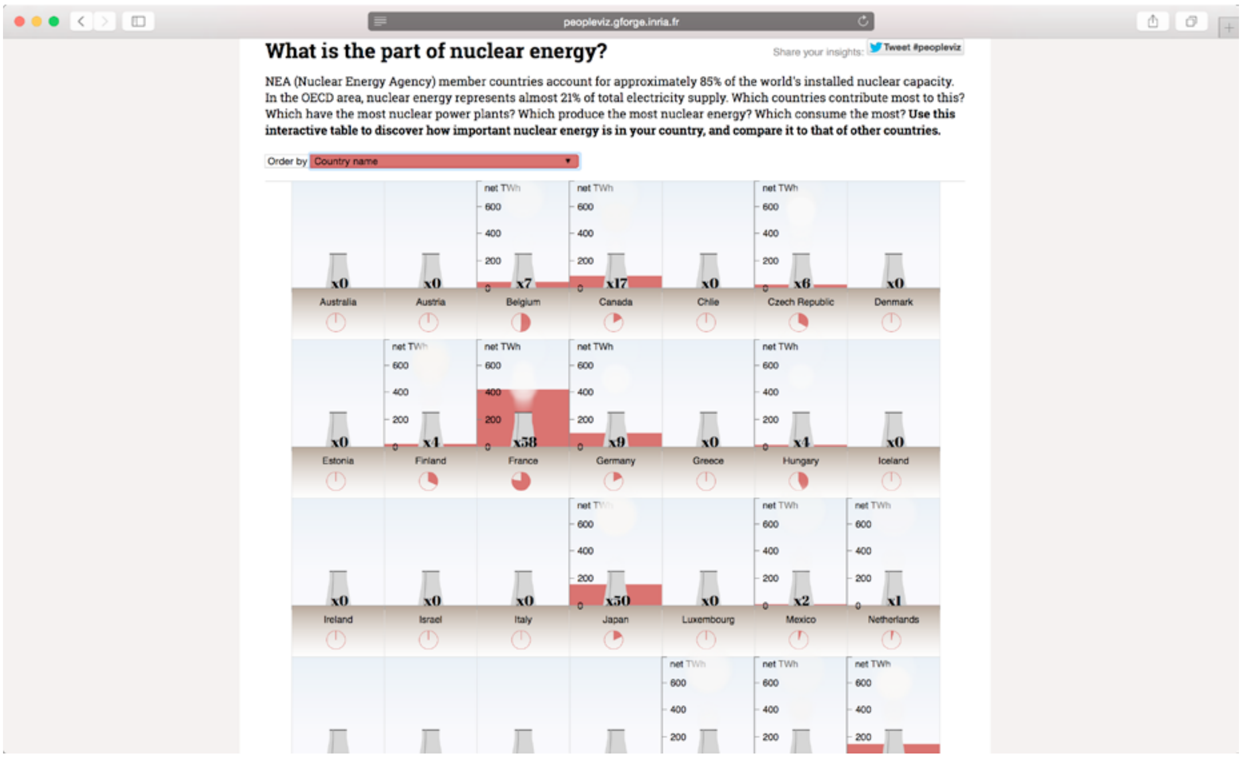


FIGURE 6.13: The *Nuclear Power Grid*
(the simpler visualization)—Explore section.

Grid [108]—for which we recreated the two alternately assigned versions (ST and no-ST), thus respecting the between-subjects experimental design. In these, we attempted to solve the listed problems, which we summarize below. For each, we give a design rationale and the solution we adopted.

P1—A minority of visitors inspected the six sections in a linear fashion. **Rationale:** People should be aware that the stepper linearly sequences the story. **Solution:** We added a mention beneath the descriptive paragraph on the first slide to tell people that each step corresponds to a section, and that they can read through sections using the stepper-buttons.

P2—The narrative component may have provided too many insights, which may have hindered visitors' incentive to explore the visualization. **Rationale:** To foster exploration, the story should serve as a means, a "jumping-off point," not as an end. **Solution:** We told the story from a specific perspective, creating a particular theme, which left more room for discovery of important insights outside of the theme.

P3—Visitors may have been unable to articulate initial questions about the data, even with the help of the narrative component. **Rationale:** People need explicit help to articulate questions if they are not familiar with the data. **Solution:** We added explicit questions in the Explore section.

P4—Visitors may not have been sufficiently 'trained' to use the interactive features. **Rationale:** The narrative component should also provide an explicit tutorial for the visualization. **Solution:** We added a bolded instruction for each new interactive feature made available in the narrative component.

P5—The data may not have been rich enough for visitors to truly engage in exploration. **Rationale:** The dataset should hold the promise of finding interesting information for people to engage in information interaction. **Solution:** We used a richer dataset for one of the new webpages, and a simpler dataset for the other to act as a baseline.

P6—Visitors may have considered that the interactive potential of the interface was too limited. **Rationale:** People should find the interface appealing, and should be able to easily distinguish and use its different interactive features. **Solution:** On the richer dataset webpage, we added several interactive features, including direct manipulation of data objects. We emphasize that these problems and design solutions are not necessarily new, nor are they standard. We simply point them out here, as we believed they might have confounded our previous results.

Like the *CO2 Pollution Explorer,* we published both new webpages first in

English on visualizing.org, then in French on **Mediapart**. The *Economic Return on Education Explorer* was soon exhibited in the “Visualizing Highlights: August 2014” on visualizing. org, and it received a total of roughly 1300 unique browser connections in one weekend. Unfortunately, the Nuclear Power Grid did not meet the same success; it received only 119 browser connections from **Mediapart**, and 131 from visualization galleries.

6.3.1 Design

The *Economic Return on Education Explorer* (which we refer to as the *richer visualization*—FIGURE 6.12) used a rich dataset on the lifetime costs and benefits of investing in different levels of education in the OECD area; its main graphical component was an interactive stacked bar chart. There were four sections in the narrative component in the ST version (see Appendix N), which followed the layout shown in FIGURE 6.3 (A); the Explore section followed a similar layout to FIGURE 6.3 (B), except it included only one graphic.

The *Nuclear Power Grid* (which we refer to as the *simpler visualization*—FIGURE 6.13) used a simple dataset on nuclear energy production and consumption in the OECD area; its main graphical component was a table. Each cell contained a numeric value, a bar chart, a pie chart, and an illustration of a cooling tower. There were three sections in the narrative component (see Appendix O), and the layouts were again the same, except that the Explore section did not include the list (FIGURE 6.3 (B)-(4)), and query-buttons (FIGURE 6.3 (B)-(3)) were replaced by a drop-down menu.

6.3.1.1 Metrics

We created the following taxonomies of semantic operations for:

The Richer Visualization—

* **inspect:** show the specifics of the data

[*hover label, hover stacked bars*];
* **filter:** show something conditionally [*click list checkbox, click "Show All Countries/Remove All Countries" button*];
* **explore:** show something else [*click query-button*];
* **reconfigure:** show a different arrangement [*click stacked bars*]; and
* **narrate:** show a different section [**click stepper-button**].

The Simpler Visualization—

* **inspect:** show the specifics of the data [*hover background bar, hover pie chart*] ;
* **reconfigure:** show a different arrangement [*select from drop-down menu*]; and
* **narrate:** show a different section [*click stepper-button*].

Since we received fewer visits for the simpler visualization, and since our previous results had shown that there was no important difference in trends between the information-savvy and the visualization-savvy populations, we aggregated the data of both populations for the two visualizations. We performed all initial filtering and coding in the exact same way as in the *CO2 Pollution Explorer* case; in the end, we kept subsets of **1178 sessions** for the richer visualization, and of **160 sessions** for the simpler visualization. While this last number is quite small compared to those of the other cases, it is still big enough for estimation of user-behavior.

6.3.2 Hypotheses

We maintained the same qualitative hypotheses as for the *CO2 Pollution Explorer* case, and thus the same quantitative hypotheses. However, the purpose of having created two new webpages was to see whether the richness of the dataset might affect the impact of the narrative component on user-engagement in exploration. Thus, we added a third qualitative hypothesis: **the impact of the narrative component on user-engagement in exploration should be more pronounced when the visualization presents a richer dataset**, resulting in higher user-activity levels in the Explore section of the richer visualization than in that of the simpler visualization.

6.3.3 Results

For both webpages, we conducted the exact same analysis as before. We began by removing all bounced sessions—**27.9%** in the richer visualization case, and **33.1%** in the other—and we plotted all results of the whole webpage and Explore section analyses [steps 3 to 5 and 8 to 10] in FIGURE 6.14. These are compared to those of the information-savvy population in the *CO2 Pollution Explorer* case.

6.3.3.1 Narrative Framework of the Richer Visualization Analysis

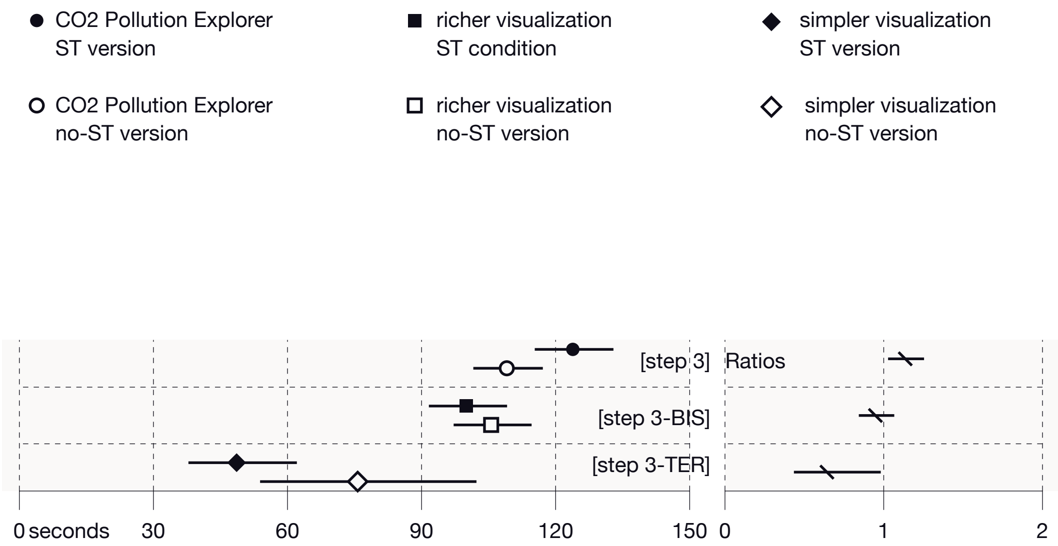
[step 6-BIS]—In **92.2%** of all sessions, visitors saw more than one section; in **57.4%**, they saw the Explore section; and in **53.7%**, they saw all five sections. The return IDs showed that **82.2%** (157/191) sessions in which visitors did not inspect all sections were first-time connections.

[step 7-BIS]—We removed these sessions from further analysis ([steps 7-BIS to 10-BIS]). In **40.9%** (91/222) remaining sessions, visitors inspected all four sections and the Explore section in a linear fashion.

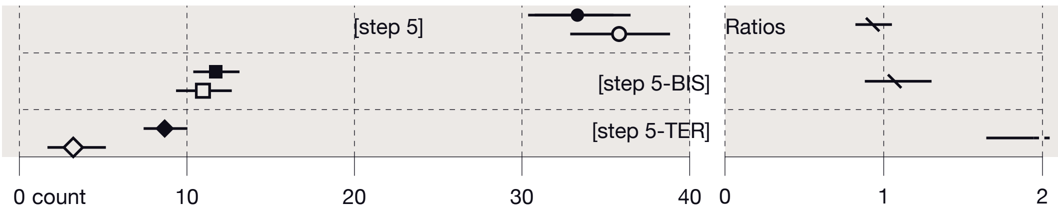
6.3.3.2 Narrative Framework of the Simpler Visualization Analysis

[step 6-TER]—In all sessions, visitors saw more than one section; in **80%**, they saw the Explore section; and in **75.7%**, they saw all four sections. The return IDs showed that **23.5%** (4/17) sessions in which visitors did not inspect all sections were first-time connections.

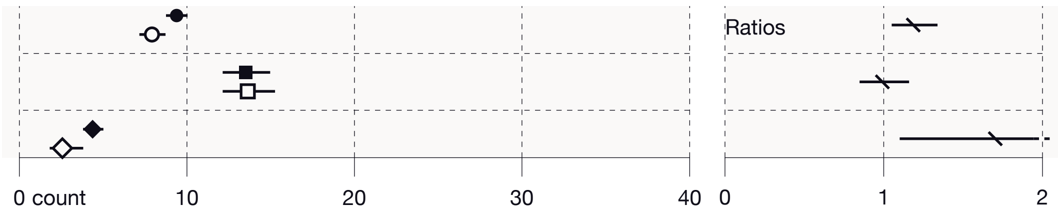
[step 7-TER]—We removed these sessions from further analysis ([steps



(A) Geometric mean uptime with 95% CI and ratio.

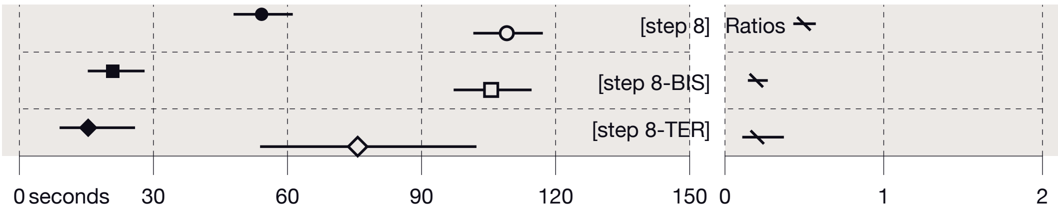


(B) Number of meaningful hover interactions with 95% CI and ratio.

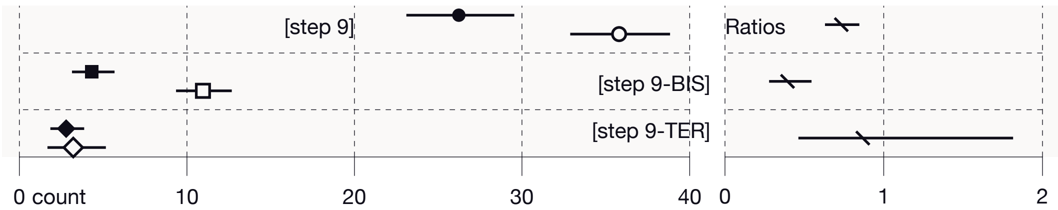


(C) Number of meaningful click interactions with 95% CI and ratio.

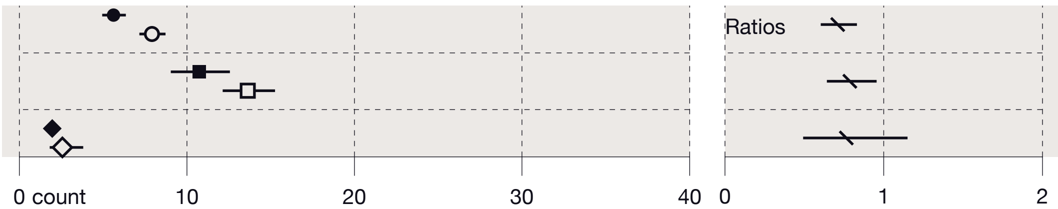
FIGURE 6.14: Activity comparisons between versions for the richer and simpler visualizations. Values for the CO_2 Pollution Explorer are also shown for reference (Figure is continued on next pages).



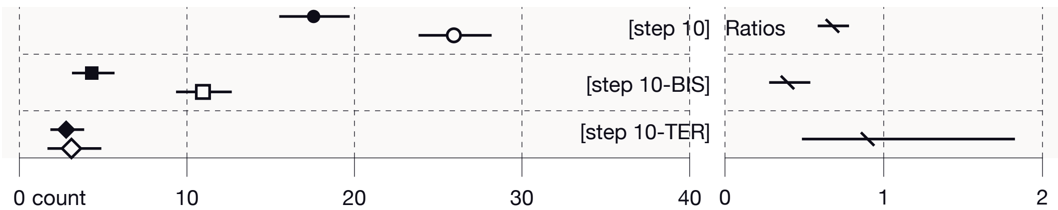
(D) Geometric mean time spent in the Explore section with 95% CI and ratio.



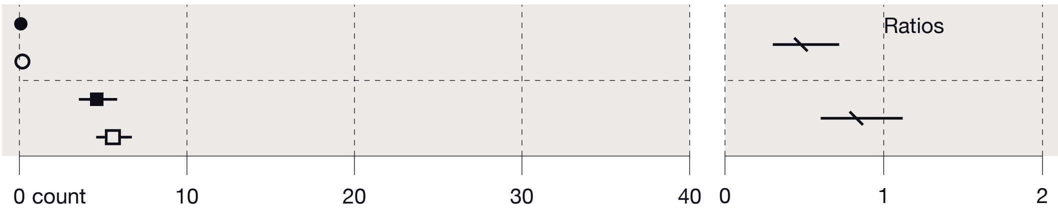
(E) Number of meaningful hover interactions in the Explore section alone with 95% CI and ratio.



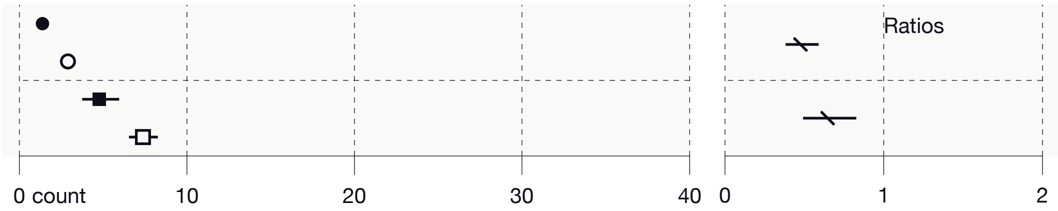
(F) Number of meaningful click interactions in the Explore section alone with 95% CI and ratio.



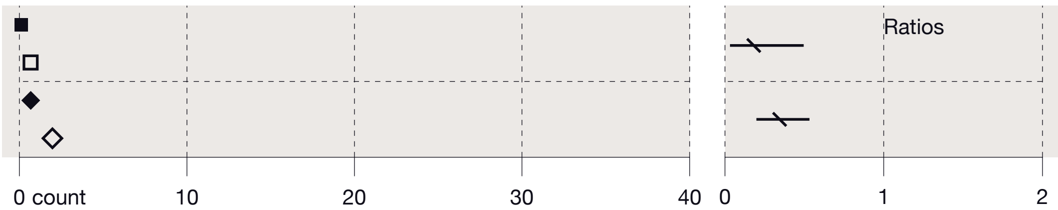
(G) Number of *inspect* operations in the Explore section alone with 95% CI and ratio.



(H) Number of *filter* operations in the Explore section alone with 95% CI and ratio.



(I) Number of *explore* operations in the Explore section alone with 95% CI and ratio.



(J) Number of *reconfigure* operations in the Explore section alone with 95% CI and ratio.

7-TER to 10-TER]). In **41.5%** (22/53) remaining sessions, visitors inspected all three sections and the Explore section in a linear fashion.

6.3.4 Discussion

In the richer visualization case, none of the 'whole webpage' and 'Explore section only' hypotheses are confirmed (**H1.1**, **H1.2**, **H3.1**, **H3.2**). In fact, there is even no evidence of a difference in total uptime, or in number of meaningful hover and click interactions between versions on the whole webpage level—as attested by the ratio 95% CI that all overlap 1 [steps 3-BIS, and 5-BIS] (FIGURE 6.14).

In the simpler visualization case, none of the 'whole webpage' and 'Explore section only' hypotheses are confirmed either, with the exception of **H1.2.** However, these results are to be considered cautiously, since they show a lot of variability in the data—as attested by the very wide 95% CI. This can be attributed to the smaller sample size. Nevertheless, since we are not directly interested in effect sizes, but rather in simply seeing if there is a difference between versions, the ratio 95% CI that do not overlap 1 (FIGURE 6.14) provide sufficient information for our needs. Overall, visitors spent the least amount of time and performed the least amount of interactions in this case, be it on the whole webpage level or in the Explore section alone—this was expected, as the dataset and interactive potential of the visualization were not as rich as in the other cases.

Finally, in both cases **H2.1** is confirmed, but **H2.2** is not. While the percentages of sessions in which visitors inspected all sections of the narrative component in a linear fashion are higher than in the *CO2 Pollution Explorer* case, they are still not a majority.

Overall, these results invalidate once again our two main qualitative hypotheses, and confirm those of the *CO2 Pollution Explorer* case: the narrative components did not immerse visitors in the way we expected them to in either cases; and they did not increase visitors' engagement in exploration in the Explore sections. Furthermore, while there is evidence that visitors of the richer visualization performed more meaningful click interactions in the Explore section than visitors of the simpler visualization did—which seems normal, since there were many more clickable features in the former—there is no evidence that they spent more time there, or that they performed more meaningful hover interactions—as shown by the 95% CI for the analysis of the Explore sections of the ST versions in [steps 8-BIS, 8-TER, 9-BIS and 9-TER] (FIGURE 6.14). Thus, there is no evidence that the narrative component in the richer visualization had a bigger effect on user-engagement in exploration than the one in the simpler visualization—this invalidates our third qualitative hypothesis. However, from a broader perspective, confirming this hypothesis would have been pointless, since each of our experiments has shown that including a narrative component does not increase user-engagement in exploration.

6.4 Conclusion

This chapter has focused on the **initial incentive cost**, and has addressed the following research question:

Q4: Can providing initial incentives for exploration, *i.e.*, external motivations, in the design of visualizations trigger an exploratory behavior in casual audiences, and lead these people to engage in efficient personal data-explorations?

I have presented three web-based filed experiments, which I conducted with Françoise Detienne and Jean-Daniel Fekete. These have shown that augmenting exploratory information visualizations with initial narrative visualization techniques and storytelling does not truly help engage users in exploration, which suggests that 'pushing' initial questions and observation as external motivations for exploration does not provide sufficient help for casual users to overcome the **initial incentive cost** (**Q4**). Nevertheless, our results are not entirely negative. The *CO2 Pollution Explorer* and the *Economic Return on Education Explorer* were 'successful' webpages that did engage people in a certain way: both received a relatively high number of visits[3], and the average uptime was well-above web standards, whatever the version. They were also curated in referential online visualization galleries, and several discussions took place around them on different external websites (*e.g.*, on Mediapart and on citylab.com). This hints to some form of social-data analysis, although most of it relied on users' background knowledge or understanding of the different topics presented in the visualizations. Thus, beyond the spectrum of this study,

3 In fact, since the time of our study the Economic Return on Education Explorer has received nearly 84, 000 visits on visualizing.org alone.

it is important that the concept of engagement in data-analysis (social or not) be better defined. Here, we consider it from a behavioral perspective as an investment in exploration, which may lead to insight. However, engagement can also be considered from an emotional perspective as part of an aesthetic experience, as is done with certain casual information visualizations [[226], [275]], or from a social perspective, in which case a visualization is only a vector for discussion and debate.

In addition, as our approach of using behavioral metrics is original, it is hard to determine whether our visualizations were generally 'engaging' or not (beyond users' propensity to explore the data). Unlike for other webpages, there is no baseline for the number of visits, the time spent on the webpage, *etc.,* for online visualizations. Although we have compared our results with metrics established for other webpages and sites, this seems somewhat inappropriate because visualizations offer a much greater potential for interaction and exploration. Thus, we believe more work should be conducted in this direction, and more data should be collected and shared to establish a common understanding of what to expect from engaging visualizations.

Finally, although our goal is to overcome online users' limited attention span, we cannot rule out the fact that our experiments may have failed simply because of this. If people read through an introductory narrative component, they may not want to spend extra time exploring a visualization. Nevertheless, we still believe that 'pushing' observations, unanswered questions, and themes from a narratorial point of view in the form of an introductory story is insufficient for engaging people to dig further for personal insights. Ultimately, we hope that this work and the data we have collected will contribute to establishing a baseline for investigating other strategies that may provide initial incentives for exploration to users.

Chapter 7

Conclusion

In this dissertation, I have highlighted several initial challenges casual audiences may encounter when confronted with online information visualizations of open data. Based on an analogy with the theory of information foraging [222], I have identified four sub-costs of van Wijk's perception and exploration costs (Ce) [263]: a **literacy cost**, a context-interpretation cost, a **perceived interactivity cost**, and an **initial incentive cost**; and I have shown how these can theoretically be articulated around the concept of engagement using O'Brien & Toms' for-stage-model (FIGURE 2.2) [214]. I have also presented the constructs behind each of these sub-costs, and have used them to review several success stories and acknowledged failures of infovis for *the* people. For each, I have set specific research questions that I have addressed assuming either an evaluation approach, or a design approach; in one case, I adopted both approaches (in Chapter 5). Coming from a graphic design background, I believe this is one of the main originalities of this work, as designers generally tend to ignore evaluation of the systems and artifacts they create, while infovis researchers often overlook aspects of design and visual communication.

In retrospect, I consider this interdisciplinary approach was what led me to focus on integrated visualizations techniques for independent use in a broad audience. Traditional infovis research generally focuses on non-integrated techniques, *i.e.*, visualization and/or interaction techniques that do not have to take a website or news article's general layout, styling, and standard interaction techniques into consideration. Granted these are extra experimental factors, which make it much harder to isolate individual variables to test—this was one of the challenges I encountered when working with Mediapart. However, working only with non-integrated techniques prevents studies from being conducted 'in the wild.' While such studies have their issues—as they generally lack control over certain experimental conditions (*e.g.*, type of display, demographics of users, *etc.*) and force the researcher to go beyond a simple prototype to a fully functional system that can work without his/her assistance—they provide more ecologically valid insight into what people do with visualizations, outside of traditional controlled environments. In addition, designing contextually integrated systems for independent use generally requires a heavy graphic and interaction design input, which takes time and effort to consider and implement, and which is usually beyond the scope of what infovis researchers want to evaluate. This is why, in my opinion, the sub-costs of Ce have often been overlooked (typically the **context-interpretation cost** and the **perceived interactivity cost**), even though addressing some of them is a standard issue for designers.

This last chapter is organized in the following way. It begins with a general summary of the research presented in this dissertation, and relates it back to the analogy with the theory of information foraging. Section 7.2 then discusses this work with regard to the general research question set in Section 1.2, and describes some of the lessons I learned using Amazon's Mechanical Turk (AMT) while conducting crowdsourced experiments. Finally, Section 7.3 details and extends several perspectives I have already mentioned for future work on two sub-costs of Ce.

7.1 General Summary

In this section, I summarize the different contributions of this dissertation, with regard to the lower-level research questions set in Section 1.2.4, and to the sub-costs of Ce these relate to. I follow the same order as in the previous chapters, *i.e.*, 1) the **literacy cost**, 2) the **context-interpretation cost**, 3) the **perceived interactivity cost**, and 4) the **initial incentive cost**. For each, I also return to the original analogy with information foraging.

7.1.1 The Literacy cost

If an information seeker is not used to 'reading' from visualizations, s/he is likely to perceive a high cost/benefit ratio in putting the cognitive effort into understanding the graphic. To address this **literacy cost**, I had set the following research question:

Q1: How can a designer know the level of understanding an audience has of different visual representations of data?

Knowing how well people understand visualizations can have a major influence on design choices. It can help determine what kind of representations to use, and if visualization is altogether an appropriate medium for a targeted audience. For example, if the audience's level of visualization literacy is too low, other media like text may be preferred. To address **Q1**, I adopted an evaluation approach. In Chapter 3, I proposed a definition for visualization literacy, and I developed a method for assessing the visualization literacy of a user, based on a principled set of considerations. I used Item Response Theory to separate out the effects of item difficulty and examinee ability, and I designed a series of fast, effective tests for line graphs, bar charts, and scatterplots. I then tailored these tests to provide immediate estimates of a user's visualization literacy, and published them online at http://peopleviz.gforge.inria.fr/ trunk/vLiteracy/home/. In addition, I have made the source code available on GitHub for versioning [101]. As such, the main contributions of Chapter 3 were as follows:

* a practical definition of visualization literacy;
* a method for: 1) assessing the relevance of visualization literacy test items, 2) assessing an examinee's level of visualization literacy, 3) creating fast and effective assessments of visualization literacy for well established visualization techniques and tasks; and
* an implementation of four online tests, based on our method.

7.1.2 The Context-interpretation cost

In the case of most online information visualizations, an information seeker must first read a series of titles, labels, annotations, *etc.* to find out what the data are about before s/he can even begin to estimate the cost/benefit ratio of exploring those data. To address this **context-interpretation cost**, I had set the following research question:

Q2: How can visualizations be designed to help people interpret their context, *i.e.*, the semantic nature of the data they present?

Although information visualization design and graphic design are related disciplines, both interested in how graphics can be organized to present information, they are quite distant. I believe there is a great opportunity for bridging the gap between them, as visualizations are now often used as media for journalism. To address **Q2**, I adopted a design approach. In Chapter 4, I described the design of the *CO2 Pollution Map,* a visualization that takes inspirations from the disciplines of graphic and motion design to convey semantic information about the data using free visual variables. This example illustrates how transposing design considerations used in other fields—specifically the interpretation level of visual communication—can impact visualization design choices. Based on this, I proposed a framework to help rethink visualization design from a visual communication perspective (FIGURE 4.4). This model shows a continuum between traditional infovis design and graphic design approaches, and emphasizes the fact that a viewer first needs help interpreting what a visualization is about (at the interpretation level). As such, the main contributions of Chapter 4 were as follows:

* a set of considerations derived from the design of the *CO2 Pollution Map;* and
* a novel framework for thinking about visualization design.

7.1.3 The Perceived interactivity cost

If an information seeker reaches a visualization patch expecting passive interaction, and if the visualization does not provide any cue as to its interactivity, s/he is likely to perceive a high cost/benefit ratio in attempting to discover if the graphic is interactive, and what interactions it provides. To address this **perceived interactivity cost**, I had set the following research question:

Q3: Do online users have a natural propensity to interact with visualizations—especially when these are embedded with text—and if not, how can we help these people detect the interactive potential of information visualizations?

Oftentimes, information visualizations are developed independently from the context in which they are published. On a webpage, these usually end up embedded with other media like text—something rarely considered by designers; visualizations are then simply additional means for communicating information. In addition, visualizations in such contexts have traditionally been static (*e.g.,* infographics), and many still are. Therefore, it is important to make sure people perceive and understand the interactive potential of information visualizations, so that they can engage in data-explorations. To address **Q3**, I adopted both an evaluation and a design approach. In Chapter 5, I first assessed that a majority of people lack initial propensity to interact with simple visualizations when these are embedded with text. I then proposed a design space for Suggested Interactivity (SI—FIGURE 5.6), and illustrated how it can be used to create cues that can help users identify the interactivity of abstract interface features like visualizations. Based on this, I evaluated three different SI cues for interactive bar charts embedded in Wikipedia pages, and found that when a cue provides feedforward, it can successfully entice more people to perform interactions. Interestingly, in relation to the **literacy cost**, I also found that people can rapidly increase their visualization literacy level when questions and charts are highly congruent. As such, the main contributions of Chapter 5 were as follows:

* an assessment of the need for SI in cases where visualizations are embedded in with text;
* a design space for SI;and
* an evaluation of three different SI cues for bar charts, which we created using specific design considerations derived from our design space.

7.1.4 The Initial incentive cost

Finally, if an information seeker does not have sufficient background knowledge about a visualized dataset or the indicators it uses, s/he may find it hard to articulate initial questions, which may generate a lack of motivation to explore the data. Similarly, if s/he expects to find information upfront, s/he is likely to perceive a high cost/benefit ratio in attempting to dig for it. To address this **initial incentive cost**, I had set the following research question:

Q4: Can providing initial incentives for exploration, *i.e.*, external motivations, in the design of visualizations trigger an exploratory behavior in casual audiences, and lead these people to engage in efficient personal data-explorations?

Narrative techniques and storytelling in information visualization have sometimes been considered a possible means for motivating people to dig for extra insights, once the main information has been communicated. This suggests that providing users with initial observations, unanswered questions, and themes may provide sufficient incentives to engage these people in data-explorations. To address **Q4**, I adopted an evaluation approach. In Chapter 6, I presented three web-based field experiments, which have shown that 'pushing' such initial questions is insufficient. However, I also stressed that our results should be considered carefully, as they only describe engagement from a behavioral point of view. This approach is original, which makes it is hard to compare to other visualizations that may be considered 'successful' or engaging, as their evaluation has generally relied on qualitative observations like comments. Overall, the three visualization websites I tested did generate a good amount of traffic, and stimulated several interesting discussions on various external websites. This means they were engaging in a certain way—although not in the one we had expected—and encourages me to keep working with log data, in order to establish a behavioral baseline for success and 'engaginess' of visualization design. As such, the main contributions of Chapter 6 were as follows:

* a first large-scale assessment of how users behave with different online information visualizations; and

* three assessments of the impact of 'pushing' observations, unanswered questions, and themes as initial incentives for exploration on users' behavior.

7.2 General Discussion

Beyond the lower-level research questions discussed in the previous section, which motivated and structured the individual projects presented in this dissertation, I have focused on addressing the following general research question:

How might the perception and exploration costs associated with using an information visualization limit people's engagement in efficient explorations of data, and how might these limitations be remedied?

In this section, I first discuss the problem of establishing who *the* people are; after that, I relate the work presented in this dissertation back to this general research question. I then finish by presenting several lessons I learned using AMT while conducting crowdsourced experiments.

7.2.1 Defining the People

One of the main challenges of this work has been to define who *the* people are, *i.e.,* the targeted audience. Identifying a clear audience is an important aspect of user-centered design, and it can help make appropriate design choices. In this dissertation, I have assumed *the* people are casual audiences, which I define as people who are usually confronted with visualizations only 'on the fly,' and who do not necessarily have a high degree of domain knowledge about given data, or of infovis systems. However, domain knowledge is truly hard to assess over the web, and I now stress how broad this definition is. The perception and exploration costs may vary greatly within these people, and I believe establishing solid design guidelines for such a population is difficult, and possibly inappropriate.

In Chapter 6, I made a distinction between information-savvy and visualization-savvy populations, based on how people reached the different visualization websites used in our experiments. This is a rough distinction, and results for the *CO2 Pollution Explorer* case (see Section 6.2.4) seem to indicate that there is no real difference in behavior between these populations. A simple explanation for this is that people who reach a visualization through an online news and opinion outlet may very well be interested in information visualization, although they may not expect to find the medium there. This would classify them as both information- and visualization-savvy. However, this does not necessarily mean they are visualization literate. Likewise, people who reach a visualization through an online infovis gallery may also be interested in the specific data used.

Usually, designers focus on relatively well identified audiences, with which they can discuss and iterate over design choices to best suit people's needs. These audiences are generally distinguished using market segmentation strategies or psychological indicators, either by designers themselves, or by marketing branches or agencies beforehand. While such an approach could prove interesting for identifying different populations within *the* people, I believe it would require specific market-research (or user-research) for each type or topic of open data. Typically, audiences interested in such data may range from political activists and policy makers (who will be particularly interested in some datasets and not in others), to lay people who simply want to have a general understanding of the society or world they live in. Within these groups, some people may be data- and/or visualization-savvy, while others may not; and some may be interested in establishing their own perspective on data, while others may simply want to have a general overview of public opinion. This forces me to consider that publishing visualizations on an online news and opinion outlet to test ways to engage people in the exploration of data may not be an optimal method—especially from a behavioral perspective. While such outlets help generate traffic, news readers generally only look for upfront information; they do not want to have to "work" for it [62]. Nevertheless, the comparison established in Chapter 6 between visualizations published on Mediapart and on visualizing.org seems to indicate that this was not a bias in our experiments.

Overall, I believe it is important to understand why and how people arrive at a visualization website. If a user comes from a link posted and discussed on a social network website, s/he may be more inclined to explore the data, as s/he may see potential for social interactions. Once again, although I have not directly explored social design in this dissertation, I believe it may help people overcome the **initial incentive cost** (as discussed in Section 2.3.2.4).

Finally, I also believe the notoriety and *posture* of a publisher can have an impact on what people expect from an information visualization. Today, when the New York Times publishes a visualization online, people can expect it to be informative of current events and potentially interactive. However, when a Human Rights Organization like the Center for Economic and Social Rights—which usually publishes textual communications—includes a visualization in an online article [80], people may not know what to expect.

This is why in Chapter 5 (Section 5.1.3) I tested the effect of Wikipedia templates on user's propensity to interact with charts. Although the results show no evidence of a difference in this specific case, I consider that from a broader, more ecologically valid perspective they may be biased. The 'audience' these experiments were conducted with was composed of paid workers, *i.e.*, Turkers, who generally have other motivations than everyday information seekers on the web, *i.e.*, to gain money.

7.2.2 Engaging Casual Audiences

With these user-population limitations in mind, I now return to the main research question (reminded above) and the associated assumption I had for this dissertation, *i.e.*, considering the different sub-costs of Ce in the design of online information visualizations will help engage casual audiences in efficient and meaningful explorations of open data. I also relate these costs back to the theoretical articulation I proposed in FIGURE 2.2.

First, concerning how the perception and exploration costs may limit people's engagement, the different studies presented in this dissertation have shown how each sub-cost of Ce can be challenging for uninformed or untrained users. The evaluation of Turkers' levels of visualization literacy described in Chapter 3 showed that some had trouble understanding even the simplest and most common types of visualizations (*i.e.*, line graphs and bar charts), which means they would undoubtably have a hard time trying to make sense of more complex or unconventional ones (*e.g.*, adjacency matrices, treemaps, *etc.*). Similarly, the second experiment on interaction propensity described in Chapter 5 showed that even when the (simple hover) interactivity of the visualizations was explicitly mentioned, and visualization was the only medium users had for finding relevant information to complete the given tasks, not all Turkers immediately put the effort into trying to extract that information—indeed, several needed three repetitions to elaborate and/or perfect their searching strategies. I interpret this as another indicator for low visualization literacy levels, which thereby confirms that overcoming the **literacy cost** is a prerequisite for engagement, since it can prevent people from even trying to look for information in a visualization, *i.e.*, from arriving at the point of engagement. If this cost is not overcome, there can be no effective engagement, *i.e.*, engagement with the content rather than simply with the 'pretty graphics.' In addition, the other two studies on interaction propensity described in Chapter 5 showed that a majority of Turkers did not instinctively seek to interact with the charts, meaning they could not extract the necessary information from the visualizations. However, most of those who discovered the interactivity switched to using the visualizations in the subsequent trials of the study.

This confirms that the **perceived interactivity cost** can prevent people from using a visualization, even when they have the necessary visualization literacy skills and are ready to do so to find information, *i.e.*, when they are at the point of engagement. By analogy, if this cost is overcome, people can then engage in the efficient use of a visualization. However, this is still dependent on whether they have clear questions in mind to trigger the exploration (which was the case for the studies on interaction propensity, since we provided those questions), and a general understanding of the context of the visualisation (which was also the case, as we provided a simulated task scenario). Indeed, the evaluation of popular narrative visualization techniques and storytelling described in Chapter 6 showed that even when initial questions and observations were 'pushed,' most people hardly overcame the **initial incentive cost** after reading though the introductory story, *i.e.,* after arriving at the point of engagement—or more precisely at the point where engagement [in exploration] could occur. In addition, while I acknowledge that more work is needed to truly assess the impact of the **context-interpretation cost**, I believe that referring back to the analogy with a book's cover indicates that this cost should also be overcome at the point of engagement. If a user is not attracted to a visualisation by some contextual element, s/he is unlikely to "look beyond the surface."

Second, concerning ways to remedy the limitations induced by the perception and exploration costs, the use of the different design frameworks proposed in this dissertation has shown encouraging initial results for helping people overcome both the **context-interpretation cost** and the **perceived interactivity cost**. In addition, while the ST versions of the visualization websites described in Chapter 6 seem to have failed to help people overcome the **initial incentive cost**, I do not consider the main assumption of this dissertation to be false. The *CO2 Pollution Explorer,* the *Economic Return on Education Explorer,* and the *Nuclear Power Grid* generated interesting discussions and debates on various other websites, which I ultimately consider a success. One of the advertised benefits of open data is to enlighten public debate, and although people did not engage in more exploration when provided with initial observations, unanswered questions, and themes, they did engage in some form of social data-analysis through conversation. This raises an important question that has not yet been considered, which is to know whether information visualization for *the* people should be considered as a tool for analysis, an interactive medium for non-linear communications between an author and an audience, or simply a social object which should act as a vector for discussions and debate. Each consideration could help further define the different ways in which people may engage with online visualizations, *e.g.,* in data-explorations (as assumed in this dissertation), in interaction with a 'storyteller,' or in social interactions.

Pursuing, while I have purposefully studied each of the sub-costs of Ce separately in this dissertation, I believe that specific design considerations can help address several (if not all) at the same time. For the design of the visualizations websites described in Chapter 6, I took special care to guide users through both the visual encodings and the different interactive features in the narrative components of the respective ST versions. By doing so, I hoped to help users overcome the **literacy cost** and the **perceived interactivity cost**. I also focused on designing for the interpretation level of visual communication (see Chapter 4), in order to address the **context-interpretation cost**.

As discussed in Section 5.1.2.3, it seems that providing people with tasks to conduct using *congruent visual representations* can help them rapidly increase their level of visualization literacy, leading them to overcome the **literacy cost**. Here, I define congruent visual representations as *visualizations for which the interpretation of the underlying data mainly relies on setting congruent questions* (see Section 3.2.3). For example, understanding a bar chart (and most other 1D visualizations) mainly relies on setting simple questions like "Which country has the highest value for an indicator," which easily translates to "Which bar is the highest?" I posit allowing users to get used to this process of question articulation with congruent charts can lead them to understand more complex and less-congruent ones.

The interactive slideshow format of the *CO2 Pollution Explorer,* the *Economic Return on Education Explorer,* and the *Nuclear Power Grid* did this by using different views to guide people through the visual encodings. Although the views did not invite users to perform particular tasks, the explanatory text, *i.e.,* the story, was written to be redundant with the visualizations, so that people could 'learn' to relate the information domain (or data domain) to the visual domain. I believe this is partly why people spent a relatively high amount of time on the webpages. However, I stress that this pedagogical approach may turn out to be annoying for more experienced users, who may simply want to 'get to the point.' To satisfy everyone would require having a better idea of what the targeted audience is, and of how high the sub-costs of Ce are for these people. This could help adjust design decision to best fit users' skills. Another solution would be to dynamically assess the importance of the sub-costs within different audiences (*e.g.,* individual users' levels of visualization literacy), in order to automatically update and adapt a visualization to meet their skills. Note that establishing such a dynamic balance between skill level and challenge could also be used to lead people into the state of Flow (see Section 2.2.4).

Overall, I do believe information visualization has great potential for engaging citizens with open data. However, it is important that designers thoroughly consider what kind of engagement is expected, *i.e.,* engagement in exploration, in non-linear communications between an author and an audience, or in social interactions, as visualizations may lead to different kinds of investments in understanding or debating these data.

7.2.3 Lessons Learned using AMT

To finish this discussion, I believe it is interesting to share some of the lessons I learned while conducting the experiments presented in Chapter 3 and Chapter 5 on AMT. While previous work has assessed several limitations and difficulties tied with using the platform (*e.g.,* [[175], [234]], I encountered other problems, which I did not find documented. In this subsection, I describe two of these issues.

In Chapter 5, I have presented a series of between subjects experiments, which required distinct participants for each condition. While conducting these experiments, I attempted to prevent Turkers from overlapping between conditions by explicitly mentioning in bold red font that each HIT was part of a broader study, and that if Turkers had already accepted a HIT that looked similar to the one at hand, they would not get paid for their work. Several Turkers took the HITs all the same, and one sent me the following email:

> *"We don't know if we've participated in any other HITs from you as most of us don't take note of the requesters or remember the surveys once completed. We have lives! Please be a responsible requester as many are and 1) figure out how to screen us out if we're not eligible and 2) tell us you'll do that. This way we don't waste our time and you don't lose valuable research!"*

Despite the unfriendly tone, I thanked this person, and sought out a way to screen participants who had already taken our HITs. Manually rejecting overlapping Turkers is a long and tedious task, which can only be done once they have completed a HIT. This is frustrating both for the requester and for the worker, who will have spent time doing the HIT for nothing. To minimize the effort and to prevent frustration, we implemented a simple technique for automatically screening Turkers beforehand using the data in a HIT's iframe header. The header contains each new Turker's ID upon acceptance of the HIT. By collecting the IDs of Turkers who had already participated in our running between-subject experiments and inserting them into an array inside the markup language provided by AMT, I was able to automatically check whether an arriving Turker had already participated one of the experiments. If so, a message was displayed explaining why the Turker could not redo the HIT instead of directing him/her to the study. I found this to be a useful and fairer alternative to blocking Turkers, as it does not affect their profile.

Finally, I also discovered that having a good reputation on social platforms like *Turkopticon* [69] is important. I received another e-mail from a Turker saying she would not do an experiment because our lab's Turkopticon score was too low. This was due to poor reviews our account had received during a previous experiment, almost four years before I conducted mine! Thus, I strongly encourage researchers to pay careful attention to what Turkers are saying about their experiments on such platforms, as this can have an impact on future studies.

7.3 Perspectives

In the Conclusion sections of Chapter 3, Chapter 4, Chapter 5, and Chapter 6 I have proposed several perspectives for future work regarding each of the sub-costs of Ce. In this section, I detail and extend those of Chapter 3 and Chapter 6. I keep the same order, and present avenues for future work on 1) the **literacy cost**, and on 2) the **initial incentive cost**. I also present some initial work I have conducted in these directions. Note that while I do believe there are interesting perspectives for future work on the other sub-costs of Ce, I simply chose not to detail them further than what I have done in the chapters that respectively address them. I then finish by describing perspectives for future work on the measurement of users' level of engagement with information visualizations.

7.3.1 On Visualization Literacy

As briefly mentioned in Section 3.7, I see several avenues for future work on visualization literacy. In this subsection, I detail two of these, which I have already started working on with different collaborators.

7.3.1.1 *A Behavioral Proxy for Visualization Illiteracy*

In Section 5.1.2.3, I hypothesized that seeing a progression in participants' use of charts could indicate their level of visualization literacy. Asserting this would prove useful, as it could provide a way to detect the visualization literacy of a user automatically, using log data alone. While the tests I presented in Chapter 3 can be practical in a controlled research environment, where participants can be administered short pre-tests, I fear they may be more complicated to deploy in a real environment; and in some cases, these tests may be impractical even in research environments, as they may prime participants with visualizations, biasing their judgement, or simply increasing their literacy level (as discussed in Section 5.1.2.3)—in essence, this is not an undesirable effect, but it may confound other results.

In the initial phases of our work on visualization literacy, I had sought to find such a simple indirect measure, *i.e.*, a proxy, for detecting subjects exhibiting a level of visualization literacy low enough to hamper their ability to understand visualizations. To establish this measure, I compared the progress of participants in two groups over five repetitions of twelve trials. Groups were categorized in accord with their background in infovis as visualization experts and non-experts. Performing a simple pairwise t-test on the time spent answering questions in the different repetitions of each trials, I discovered a consistent p-value for participants in the non-experts group ($p < 0.3$) when comparing the first and last repetitions, and a p-value always close to 1 in the experts group. I then conducted the same study on AMT, and found very similar results, *i.e.*, clusters of p-values below 0.3 and around 1. Although this use of t-tests and p-values is unconventional, it is not meant for significance testing but for finding a *similarity measure*. I argue that this approach is valid, as pairwise t-tests essentially boil down to comparing the average divided by the square root of the standard deviation; the test simply facilitates the application of the formula. However, I was confronted with the problem of external validity, as there were no other tests or measures to relate this to. Now that I have developed such tests, I believe this work should be extended by simply comparing this p-value threshold with the ability scores delivered by the visualization literacy tests described in Chapter 3.

Overall, developing and refining such indirect measures can have great value for the design of online information visualizations of open data, as they would allow to automatically detect a user's level of visualization literacy, and could help readjust the visual representation dynamically to fit the user's ability, thus helping him/her progressively overcome the **literacy cost** (as discussed in Section 7.2.2).

7.3.1.2 *A Framework for Research in Visualization Literacy*

In a more long-term perspective, I have started working with Sung-Hee

Kim, Sukwon Lee, Ji Soo Yi, and Niklas Elmqvist on an open visualization literacy testing platform [70]. Our goal is to provide a tool for infovis researchers to create and share tests and questionnaires that can help assess a person's level of visualization literacy. The platform is designed to help create visualizations or import already existing ones, to upload data, to create specific test questions for these, and to aggregate questions into full tests. The purpose of these tests is then to enable assessment, evaluation, and improvement of peoples' visualization literacy.

To bootstrap the platform, we held a workshop at the IEEE VIS'14 conference, in which we organized a series of *hands-on* activities. First, we invited participants to create a series of questions regarding the visualization(s) of their choice. We then instructed them to aggregate these questions (or test-items) into full-length tests, and to publish them on the platform, so that they could be shared with—and taken by—the rest of the group. This hands-on approach raised several interesting questions and comments that have allowed us to elaborate the framework for research into visualization literacy presented in FIGURE 7.1.

During the workshop, the idea of being able to read and write visualizations came up several times. This relates to the basic definition of textual literacy, and these *reading* and *writing* (or *designing*) dimensions set the first main axis of our framework. In addition, as participants in the workshop were mainly infovis researchers, the ability to assess someone's visualization literacy, and to find ways to teach it came up. These *assessing* and *teaching* dimensions set the second main axis of our framework. These two axes create four distinct spaces for research into visualization literacy.

In FIGURE 7.1, I have plotted existing work and previous events that have focused on these different spaces. First, in the top-left corner, assessing how well people 'read' visualizations was discussed during our workshop and explored using our visualization literacy platform; also in this space is the published work presented in Chapter 3 of this dissertation. Second, in the bottom-left corner, teaching people to 'read' visualizations was discussed during the EuroVis'14 conference. Third, in the bottom-right corner, teaching people to design visualizations was addressed

FIGURE 7.1: A framework for research into visualization literacy.

by Huron *et al.* [180], and explored during a workshop I co-organized with Samuel Huron, Jean-Daniel Fekete, Mathieu Le Goc, and Romain Di Vozzo. Finally fourth, in the top-right corner, assessing how well people design visualizations is, to the best of my knowledge, completely unexplored. A immediate way to populate this space could be for the infovis community to develop a visualization critique practice, as called for by Heer [174].

Overall, this framework shows how little visualization literacy has been studied. I see great potential for future work in each of the spaces, especially on teaching people to read and design visualizations.

7.3.2 On Initial Incentives for Exploration

In Chapter 6, I have shown that 'pushing' initial observations, unanswered questions, and themes from an editorial point of view does not increase users' engagement in the exploration of data. However, as discussed in Section 7.2, I believe that a more social approach can be an interesting alternative to storytelling for helping people overcome the **initial incentive cost**. I see potential in 'pulling' questions from the crowd, *i.e.,* in crowdsourcing initial incentives for exploration. Integrating social design features to online information visualizations of open data can lead people to engage in social interactions, which may encourage them to try to understand the provided data for and amongst themselves. Expert and casual audiences can exchange, debate, and challenge each other to find new perspectives. While this is similar to what Heer *et al.* have done for the *Sense.us* website [173], I believe it is interesting to see how such an approach can influence what people *do* with visualizations, *i.e.,* on a behavioral level.

As an initial step in this direction, I have implemented the *French Deputies Explorer* (Deputeviz—FIGURE 7.2) [82], a visualization that presents data on the lives and activities of French members of parliament (or *deputies*). Similarly to the visualization websites described in Chapter 6, this visualization initially guides users' through different *congruent visual representations* to help them get used to the mappings, and to help them overcome the **literacy cost**. It then opens up to *less-congruent* and more exploratory representations (see Appendix P). Essentially, the visualization is a re-organizable bubble chart that transforms into a scatterplot. Each bubble corresponds to an individual deputy, which I believe can help overcome the **context-interpretation cost**—although a proper evaluation would be necessary to truly assert this. The visualization also provides several suggested interactivity cues (*e.g.,* FIGURE 7.3) to help users overcome the **perceived interactivity cost** by showing what interactions are possible. I have integrated a *question-comment-and-response* system to the website (FIGURE 7.4), so that people can pose their own questions, and post responses to others'. Questions, comments, and responses can also

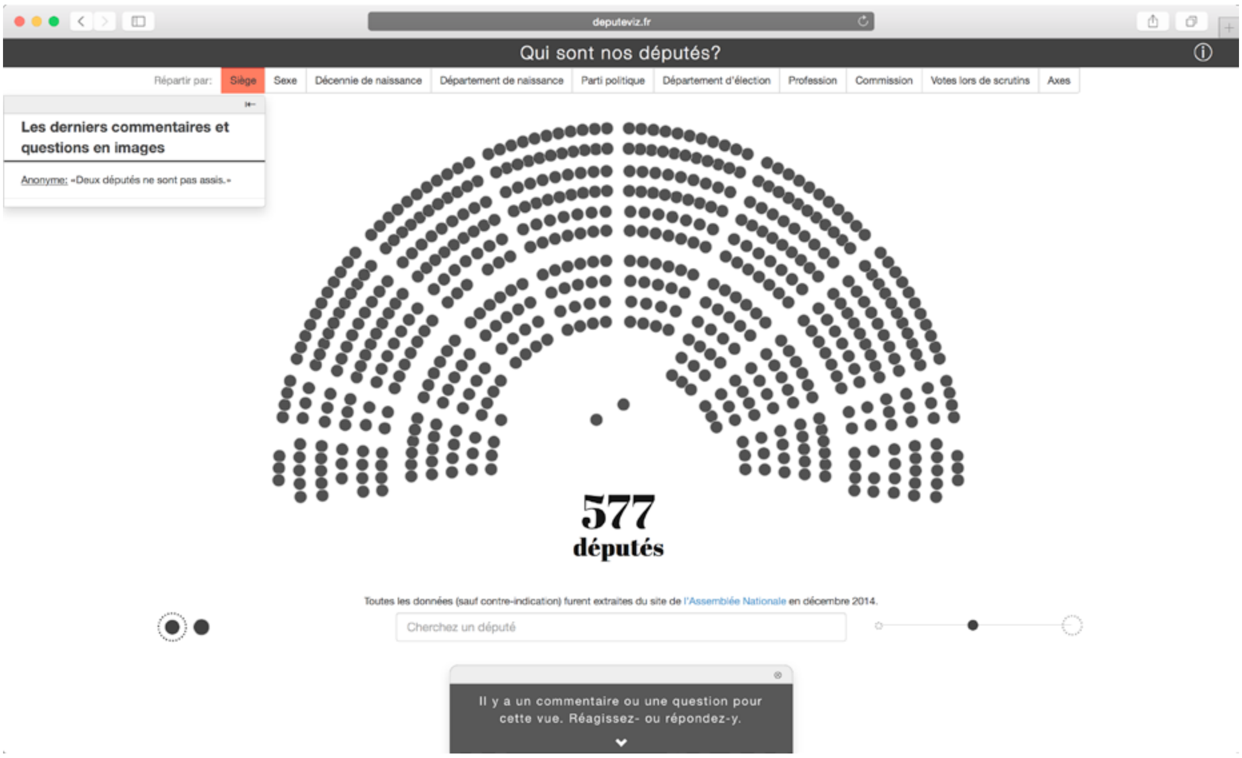


FIGURE 7.2: The French Deputies Explorer (Deputeviz).

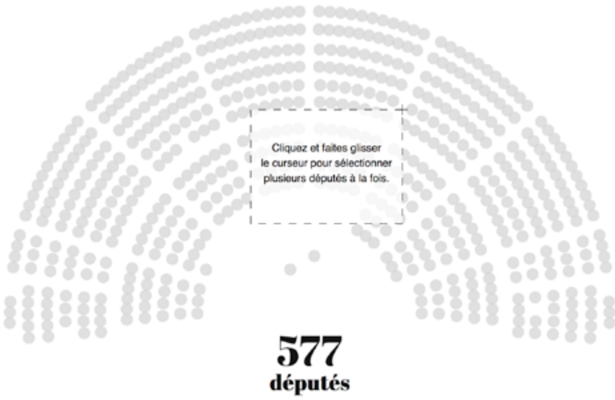


FIGURE 7.3: The French Deputies Explorer (Deputeviz)—example of SI.

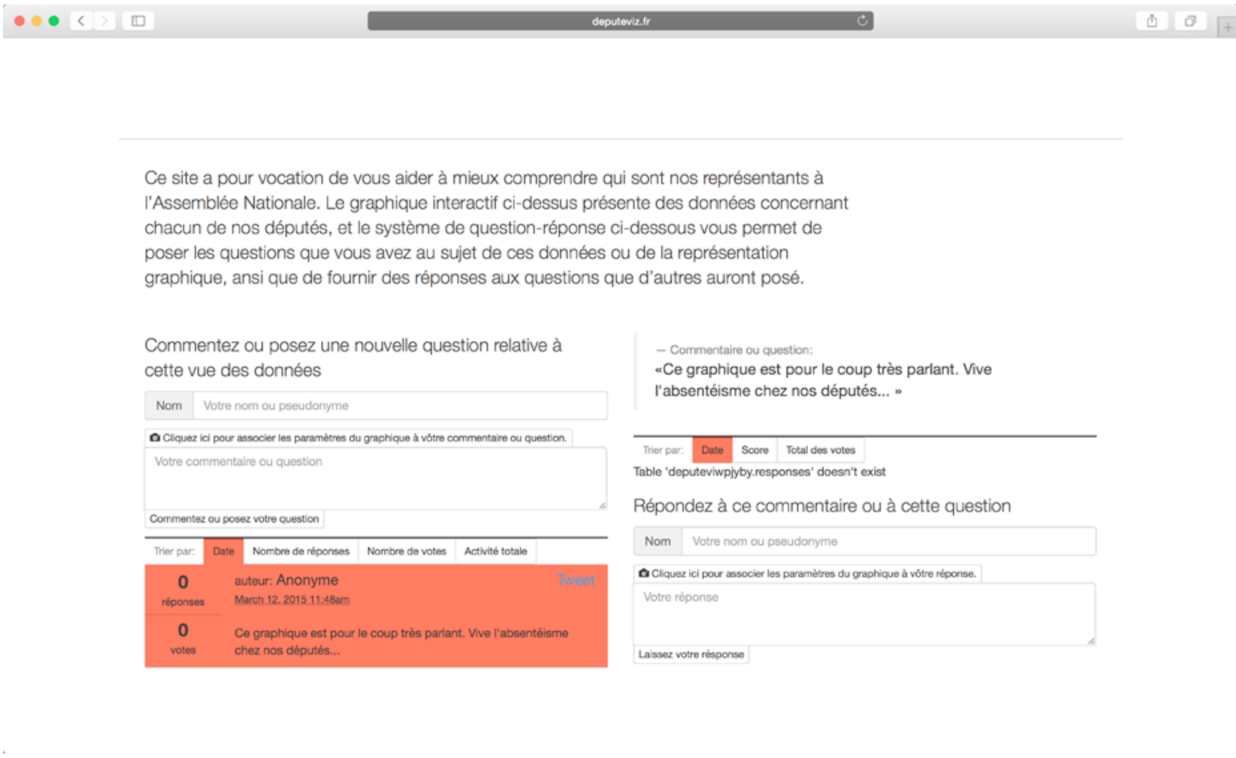


FIGURE 7.4: The French Deputies Explorer (Deputeviz)—
the question-comment-and-response system.

include bookmarked states of the visualization to help create common ground (see Section 2.3.2.4).

In an attempt to bootstrap a first set of questions, I have begun exchanging with the people at **Regards Citoyens** [96], a French online community that is interested in open political data. This has raised several interesting perspectives for the social design dimension of infovis for *the* people, which I consider as more long-term. The first perspective concerns establishing a deeper understanding of what social interactions may occur around visualizations and open data. I believe there is interesting potential here for studying the social process of collaborative sense-making using information visualizations. The second perspective concerns designing visualization interfaces that integrate social components. As it is, the question-comment-and-response system of the *French Deputies Explorer* is simply displayed below the visualization, which makes it appear below the page-fold on low-resolution screen. This can be problematic, as it separates the visualization from the possible discussions and debate. An alternative option would have been to place the social component aside the visualization (as was done for *Sense.us*), but this would have reduced the available screen-estate. I believe there is interesting potential here for exploring new interface designs that can enable social interactions. Finally, the third perspective concerns finding ways to encourage people to contribute questions, comments, and responses on the website itself, rather than discussing the visualization on other social platforms, to which they may be more accustomed to. I believe an interesting approach to this problem can be to continue exploring the use game mechanics in infovis, as done by Diakopoulos *et al.* [150]. While their simple quiz-like mechanic is interesting, it requires a heavy authoring, which seems inappropriate for the perspective of 'pulling' questions from the crowd. I believe there is interesting potential here for exploring the combination of social data-analysis and game-y graphics.

7.3.3 On Measuring Engagement

The originality of the approach described in Chapter 6 was to assess user-engagement using log data. Beyond assessing how a visualization triggers social data-analysis, I believe it is important to understand what users *do* when confronted with online information visualizations. While it is interesting to watch and analyze threads of discussions, these do not necessarily reveal whether users are actually talking about information they found in the visualization, or about their background knowledge. Our visualizations generated interesting discussions and debates on various 'external' websites, but we realized that a lot of these were not directly related to the data we used. They focused more on questions people had about the different topics our visualizations addressed. This may ultimately be a desirable goal for information visualizations of open data, *i.e.*, to be a social object for triggering public debate, but it seems unfortunate that the data are actually unhelpful for feeding the debate.

To properly measure engagement, I posit the importance of assessing how people behave with online information visualizations. This requires having a baseline understanding of what users *do* with online visualizations. Unfortunately, such metrics are rarely reported or discussed. Although in Chapter 6 we (my collaborators and I) have compared our findings with those of standard websites, I argue that may be somewhat inappropriate, as visualizations are highly interactive and dynamic media (unlike other online media).

As a short-term perspective, I will continue to collect user-traces on a large scale, in order to contribute to establishing common ground for considering engagement with visualizations. While online editors and content providers may find sufficient value in standard web-metrics like uptime and click-count, I posit these do not say much about what actually happens on a webpage. Finer analyses can be made, and I believe it should prove interesting to pursue Gotz & Wen's work on finding relations between analytic tasks and behavioral measures [166]. Likewise, it should prove interesting to find ways to combine large-scale behavioral measures with more qualitative ones (*e.g.*, comments).

As a more long-term perspective, I believe there is interesting research to be done in creating a standard tracing system that meets the specific requirements of logging users' behavior with online information visualizations. I have already started work in this direction with Jean-Daniel Fekete, as we have developed our own custom tracing system for our experiments. One of the first observations I have is that the syntax of traces should be adapted to the type of study being conducted. So far, I have mainly focused on controlled studies like the ones we conducted on AMT (in Chapter 3 and Chapter 5), and on more open A/B studies, like the ones we conducted on Mediapart and on visualizing.org (in Chapter 6). However, I believe there are other types of studies that may require different trace-syntaxes.

7.4 Final Words

"I think the idea of visualization for the masses is a good one, but not if it's also done by the masses. [...] We tend to forget all those things that now get a 'citizen' prefix (journalism, science, etc.) first had to be figured out and developed by people who spent their entire lives doing them. [...] Visualization tools may be readily available (though not all are created equal), but the knowledge how to use them well is not."—**Robert Kosara** [59]

To conclude, this work has reinforced my original belief that it is most important, as a first step, to make sure people understand the purpose and know how to use information visualizations to make sense of data before assisting them in their creation. I consider this to be part of the much border challenge of assisting the learning process of new interactive media that can help citizens better understand today's information sources (*e.g.,* data). It is of course always possible to disregard this problem at present, and to wait another ten to fifteen years for these media to become mainstream. This has been the case for computer and programming skills; it has taken roughly thirty years for public authorities—at least in France—to understanding that having a workforce and a people equipped with these skills can increase productivity and national competitiveness. However, I believe that waiting such a long time for people to learn the 'skills to be informed' is not satisfying socially, politically, or humanly. In addition, learning to use media like information visualization is undoubtably much easier than learning programming skills, so I posit it should be encouraged right now.

Overall, the purpose of the work presented in this dissertation has been to explore the initial challenges people may face when confronted with information visualizations, and to find ways to accelerate their learning using different approaches. I believe this work is important, and it should be continued. Ultimately, I hope this dissertation will contribute to develop a better understanding of how addressing these initial challenges by design can help citizens engage with the true potential of open data.

Appendix A

Open Licensing

A *license* (*i.e.,* the legal conditions under which the work is made available) is open if: 1) it allows "free use of the licensed work;" 2) it allows "redistribution of the licensed work, including sale, whether on its own or as a part of a collection made from works from different source;" 3) it allows the creation of derivatives of the licensed work and [...] the distribution of such derivatives under the same terms of the original licensed work;" 4) it allows "any part of the work to be freely used, distributed, or modified separately from any other part of the work or from any collection of works in which it was originally distributed;" 5) it allows "the licensed work to be distributed along with other distinct works without placing restrictions on these other works;" 6) it does not "discriminate against any person or group;" 7) "The rights attached to the work [...] apply to all to whom it is redistributed without the need to agree to any additional legal terms;" 8) it allows "use, redistribution, modification, and compilation for any purpose;" and 9) it does not "impose any fee arrangement, royalty, or other compensation or monetary remuneration as part of its conditions."

Nevertheless, an open license may: 1) "equire distributions of the work to include attribution of contributors, rights holders, sponsors and creators as long as any such prescriptions are not onerous;" 2) "require that modified versions of a licensed work carry a different name or version number from the original work or otherwise indicate what changes have been made;" 3) "require copies or derivatives of a licensed work to remain under a license the same as or similar to the original;" 4) "require retention of copyright notices and identification of the license;" 5) "require modified works to be made available in a form preferred for further modification;" 6) "prohibit distribution of the work in a manner where technical measures impose restrictions on the exercise of otherwise allowed rights;" and 7) "require modifiers to grant the public additional permissions (for example, patent licenses) as required for exercise of the rights allowed by the license."

Appendix B

Friel *et al.'s* Taxonomy of Skills Required for Answering Questions at Each Level

The following table is copied from [162].

Author	**Elementary** (extract information from the data)	**Intermediate** (find relationships in the data)	**Overall** (move beyond the data)
Bertin	Extraction of elementary information (*e.g.*, "What was the value of Stock X on June 15?")	Reduction in the number of data categories through combining and compiling data to discover or create fewer categories (*e.g.*, "Over the first five days, how did the value of Stock X change?")	Reduction of all the data to a single statement or relationship about the data (*e.g.*, "For the period of June 15 to June 30, what was the trend for the value of Stock X?")
Curcio	(Reading the data) Lifting information from the graph to answer explicit questions for which the obvious answer is in the graph (*e.g.*, "How many boxes of raisins have 30 raisins in them?")	(Reading between the data) Interpretation and integration of information that is presented in a graph—the reader completes at least one step of logical or pragmatic inferring to get from the question to the answer (*e.g.*, "How many boxes of raisins have more than 34 raisins in them?")	(Reading beyond the data) Extending, predicting, or inferring from the representation to answer questions—the reader gives an answer that requires prior knowledge about a question that is related to the graph (*e.g.*, "If students opened one more box of raisins, how many raisins might they expect to find?")

McKnight	Observing single facts and relationships in graphically presented data or interpreting relationships when responses involve paraphrasing the facts (*e.g.*, "What is the projected food production in 1985 for the developed countries?")	* Observing relationship within graphs an interpreting graphs as visual displays without reference to the meaning of graphical elements in context (*e.g.*, "Considering the two curves of the graph only as marks on a piece of paper, how do the changes in these two curves compare?") * Interpreting relationships either by stating that a relationship exists without describing the relationship or by making straightforward statements or relationships	* Interpreting relationships when responses require making statements that go beyond the statement of relationships to draw inferences or to recast interpretations in more technical terms. * Determining values of the data conveyed in the graph as evidence to support or reject a proposition (*e.g.*, "If this graph was offered as a piece of evidence to prove true the statement 'Storks bring babies,' how would you describe the connection between the graph and the attempt to prove the statement true?") * Assessing one's own evaluation of evidence provided by quantitative data.

Wainer	Data extraction (*e.g.*, "What was petroleum use in 1980?")	Identification of trends seen in parts of the data (*e.g.*, "Between 1970 and 1985, how has the use of petroleum changed?")	Understanding of the deep structure of the data in their totality, usually through comparing trends and seeing groups (*e.g.*, "Which fuel is predicted to show the most drastic increase in use?" or "Which fuels show the same pattern of growth?")
Carswell	Point reading or attention to a single specifier (*e.g.*, "What is the value of [the pie-slice] B?")	Local or global visual comparison of actual graph features and attention to more than a single specifier (*e.g.*, "Is [the pie-slice] A + pie slice B equal to [the pie-slice] C + pie-slice D?")	Synthesis of integration of most or all the graphed values (*e.g.*, "Is the variability of the data points large?")

Appendix C

Hartson's Summary of Affordance Types

The following table is copied from [169].

Affordance type	**Description**	**Example**
Cognitive affordance	Design feature that helps users in knowing something.	A button label that helps users know what will happen if they click on it.
Physical affordance	Design feature that helps users in doing a physical action in the interface.	A button that is large enough so that users can click on it accurately.
Sensory affordance	Design feature that helps users sense something.	A label font size large enough to read easily.
Functional affordance	Design feature that helps users accomplish work (i.e., the usefulness of a system function).	The internal system ability to sort a series of numbers (invoked by users clicking on the Sort button).

Appendix D

Jennings' Prescriptive Aesthetic Framework

The following table is copied from [184].

Unity: The coherence and completeness of objects or ideas. Unity is the wholeness of an experience.				
Techniques				
Context	*Story*	*Metaphor*	*Mini Gestalt*	*Media*
Provides richness and depth to content by creating experience.	Sets a scene that creates empathetic connection.	Assembles all parts into a whole.	Provides complete content for the particular context.	Engages all of the senses with high-quality media, particularly sound. All pieces of the environment must be harmonious (they must all fit the theme).
Justification				
Helps create cognitive continuity and memorable experiences.	Personalizes the content.	Links information for recall.	Provides a holistic and realistic portion of content.	Provides a visceral, holistic environment. The harmony and fitness of the theme is the essence of unity.

Focused Attention or Object Directedness: Elements that bring about focus or a desire to proceed with an activity.			
Techniques			
Familiarity	*Props*	*Overview*	*Media*
Go from the known to the unknown.	Use props and interactive processes.	Provide users with the big picture.	Provide interactive processes. Use eye-catching graphics and high-quality audio.
Justification			
Users pay more attention to what they know and understand.	Props help users become and remain actively involved.	An overview provides the users with a focus on which to concentrate.	Interactivity helps keep users motivated. Color and sound direct the user's attention.

Active Discovery: The process of actively seeking answers or resolutions to cognitive challenges.				
Techniques				
Problem-solving	*Play*	*Replay*	*Fill-in-the-blanks*	*Media*
Provide contextually accurate, meaningful, and purposeful guided activities.	Provide an environment that allows users to explore and experiment.	Provide users with many alternatives and options to pursue; allow them to try different things; and if they fail, provide useful feedback.	Allow users the opportunity to make connections and inferences.	When appropriate, use animation to portray concepts.
Justification				
Solving problems is a natural way to learn.	Play is a natural and active means of learning.	Once a user has succeeded at a task, s/he becomes bored and moves on. Providing feedback helps users understand where they went wrong, which provides them with information on how to be successful.	This helps users feel active and involved. It helps them feel 'smart.'	Animation can often graphically depict a procedure or concept better than words or still images.

Affect: An emotional investment that helps create a personal link to an experience or activity.			
Techniques			
Shared Experience	*First Person*	*Intrigue*	*Media*
Provide a means for users to interact with others.	Set up the environment so that users have some control over or are part of the environment.	Let the plot thicken, the story unfold, the picture evolve.	Use contextually accurate media for continuity and accuracy. Use enticing graphics.
Justification			
We, as a species, are a "small group animal," and feel a need to communicate.	The user has more of an emotional investment when the outcome is dependent on his/her input.	Surprise and mystery entices users to stay tuned-in and to remain motivated.	Users notice inappropriate media use. It breaks conversation and involvement. Enticing graphics create a desire to explore.

Intrinsic Gratification: A feeling of pleasure, reward and satisfaction from an activity.		
Techniques		
Personal Motivation	*Ownership-Investment*	*Satisfaction*
Users often come to the environment with personal motivation. Include innovative techniques to sustain motivation.	Give guided control to the user. Provide users with meaningful activities and opportunities.	Provide a means for the user to be successful.
Justification		
A user's initial desire to interact with an environment is important for motivation. Some newness, such as challenges, is refreshing.	Guidance helps users understand the environment and the content. Users want their actions to have value.	Success feels good and helps motivate users to continue.

Appendix E

Visualization Literacy: Line Graphs Test 1

The full test is available at http://peopleviz.gforge.inria.fr/trunk/vLiteracy/home/tests/lg1/.

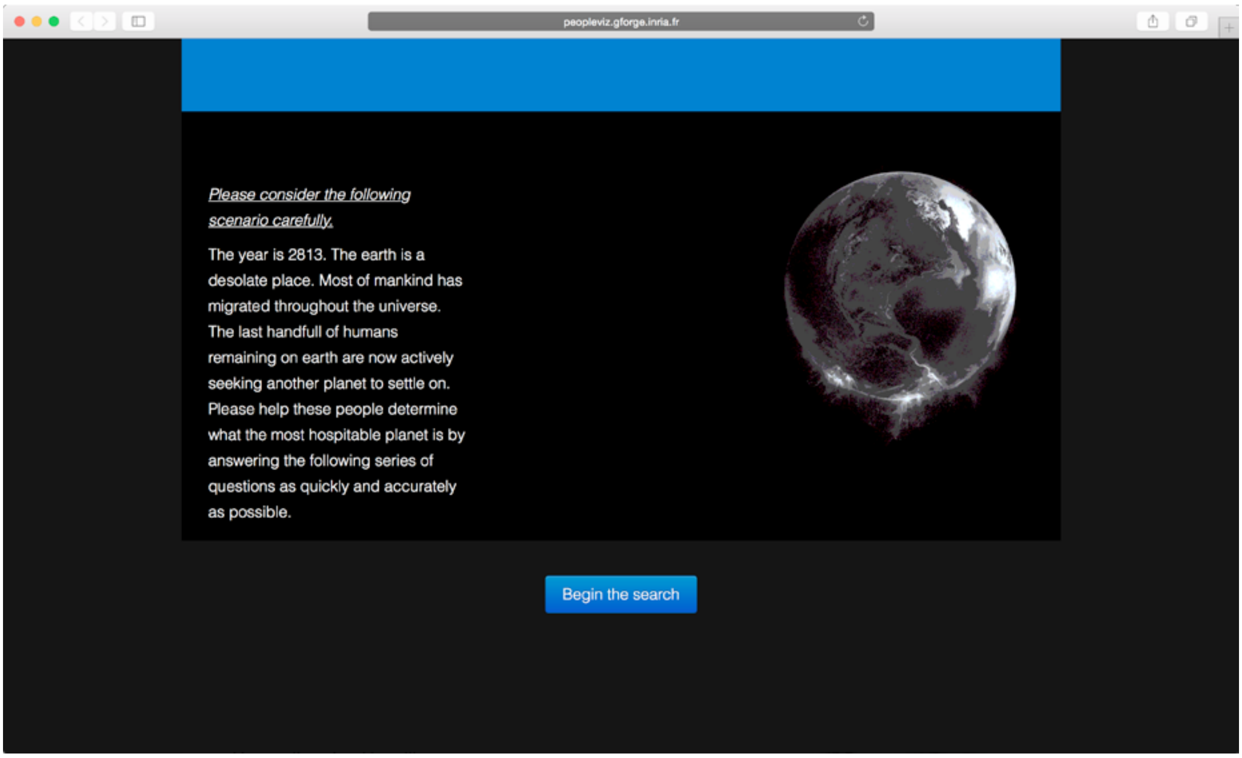


FIGURE E.1: The 9 different test-items in LG1.

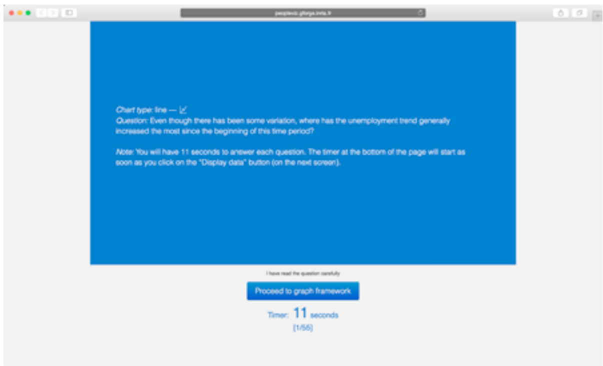
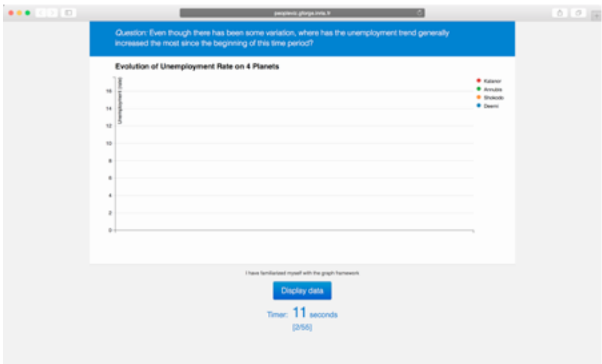


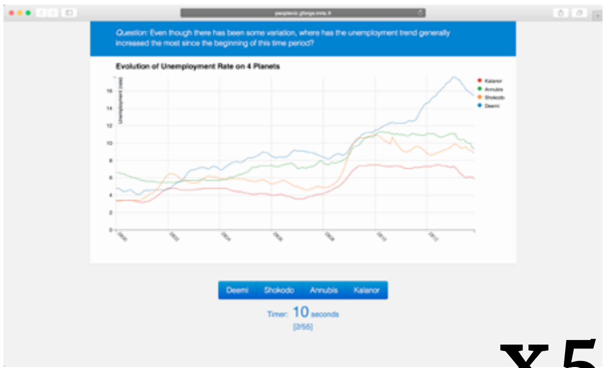
Chart type: line —
Question: Even though there has been some variation, where has the unemployment trend generally increased the most since the beginning of this time period?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[1/55]


Question: Even though there has been some variation, where has the unemployment trend generally increased the most since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[1/55]

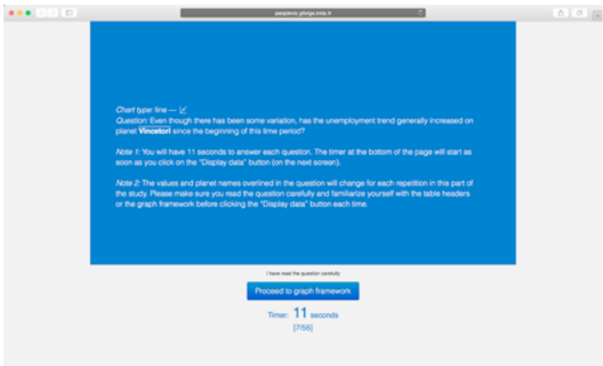

Question: Even though there has been some variation, where has the unemployment trend generally increased the most since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Krypton
Comagrave
Braal
Arcadia
Timer: 10 seconds
[1/55]

x1

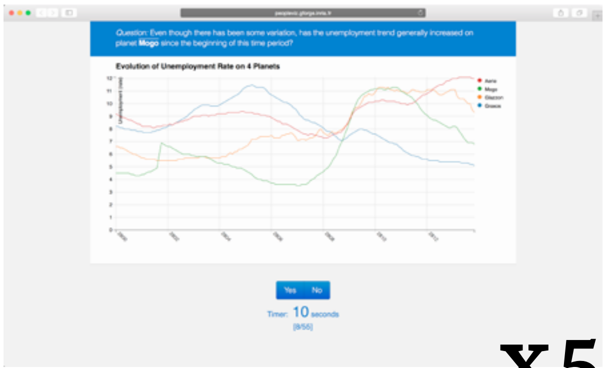

Question: Even though there has been some variation, where has the unemployment trend generally increased the most since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[2/55]


Question: Even though there has been some variation, where has the unemployment trend generally increased the most since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Deemi
Shokodo
Annubis
Kalanor
Timer: 10 seconds
[2/55]

x5



x1



x5

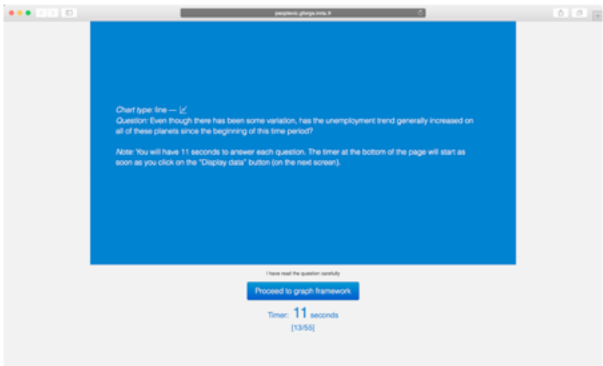


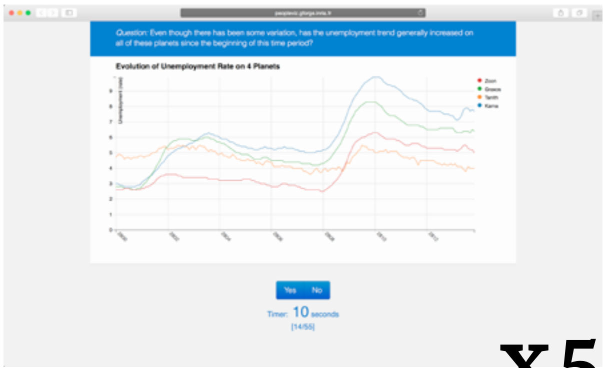
Chart type: line
Question: Even though there has been some variation, has the unemployment trend generally increased on all of these planets since the beginning of this time period?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[13/55]


Question: Even though there has been some variation, has the unemployment trend generally increased on all of these planets since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[13/55]

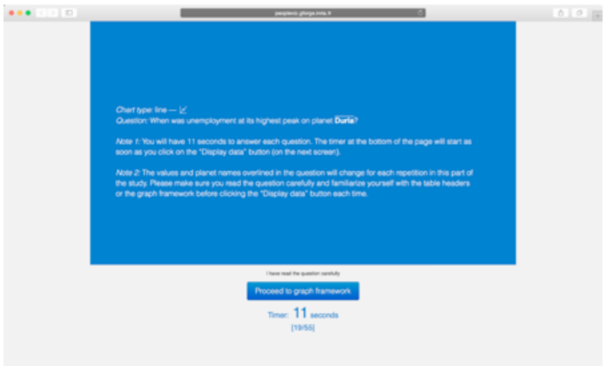

Question: Even though there has been some variation, has the unemployment trend generally increased on all of these planets since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Yes
No
Timer: 9 seconds
[13/55]

x1


Question: Even though there has been some variation, has the unemployment trend generally increased on all of these planets since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[14/55]


Question: Even though there has been some variation, has the unemployment trend generally increased on all of these planets since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Yes
No
Timer: 10 seconds
[14/55]

x5

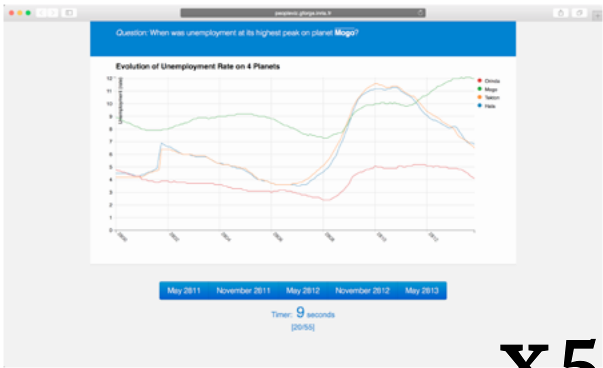

Proceed to graph framework
Timer: 11 seconds


Display data
Timer: 11 seconds

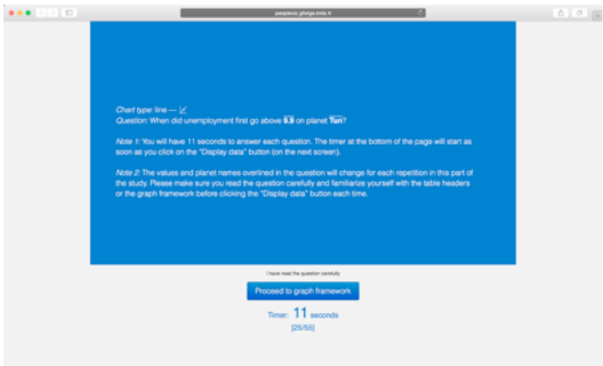

Timer: 9 seconds

x1

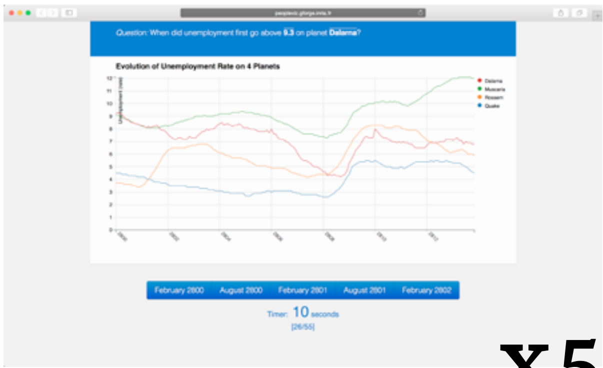

Display data
Timer: 11 seconds


Timer: 9 seconds

x5



x1



x5

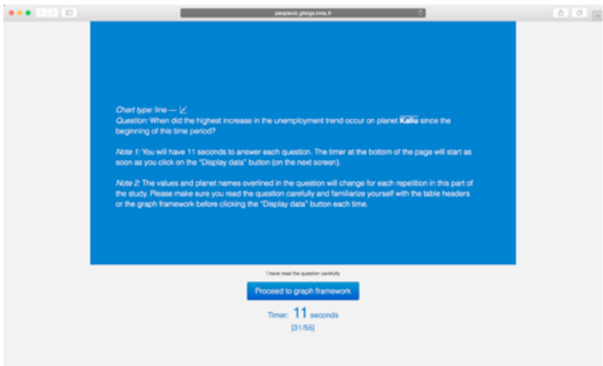


Chart type: line —
Question: When did the highest increase in the unemployment trend occur on planet Kallu since the beginning of this time period?
Note 1: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Note 2: The values and planet names overlined in the question will change for each repetition in this part of the study. Please make sure you read the question carefully and familiarize yourself with the table headers or the graph framework before clicking the "Display data" button each time.
Proceed to graph framework
Timer: 11 seconds
[31/55]


Question: When did the highest increase in the unemployment trend occur on planet Kallu since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[31/55]


Question: When did the highest increase in the unemployment trend occur on planet Kallu since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
February 2001 - June 2003
June 2003 - December 2006
December 2006 - June 2010
June 2010 - November 2013
Timer: 10 seconds
[31/55]

x1


Question: When did the highest increase in the unemployment trend occur on planet Nurag since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[32/55]


Question: When did the highest increase in the unemployment trend occur on planet Nurag since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
January 2000 - May 2003
May 2003 - November 2006
November 2006 - May 2010
May 2010 - November 2013
Timer: 10 seconds
[32/55]

x5

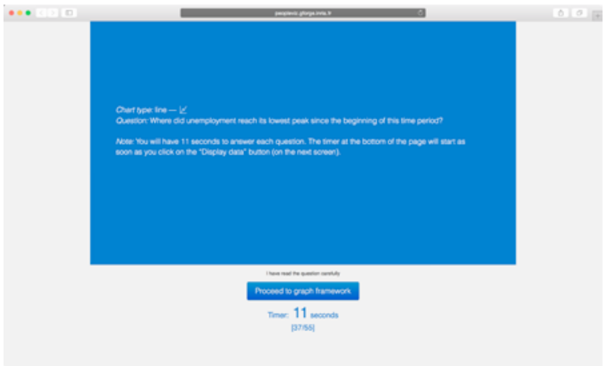
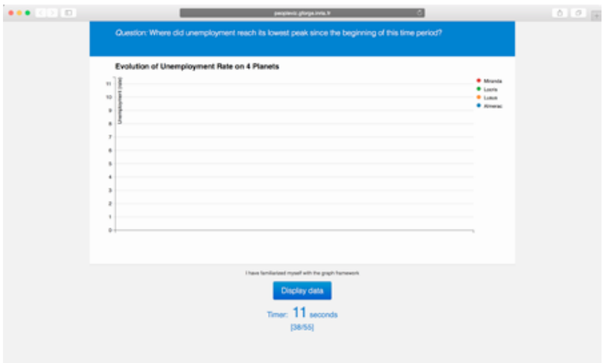


Chart type: line
Question: Where did unemployment reach its lowest peak since the beginning of this time period?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[37/55]


Question: Where did unemployment reach its lowest peak since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[37/55]


Question: Where did unemployment reach its lowest peak since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Garon
Brasl
Dysis
Darlbow
Timer: 10 seconds
[37/55]

x1


Question: Where did unemployment reach its lowest peak since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[38/55]


Question: Where did unemployment reach its lowest peak since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Timer: 10 seconds
[38/55]

x5

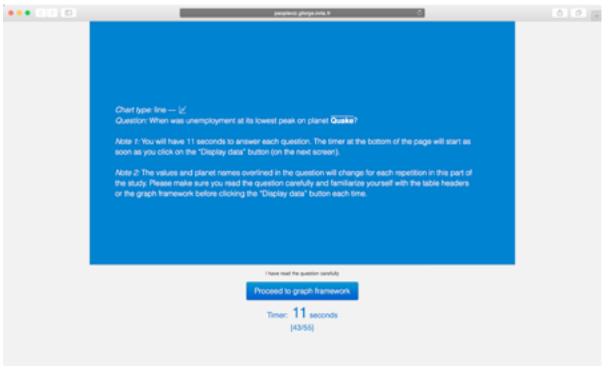


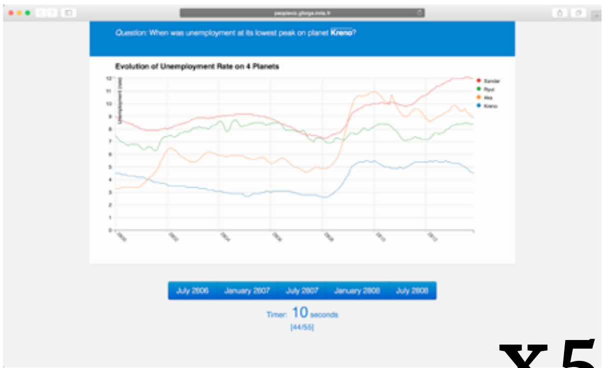
Chart type: line —
Question: When was unemployment at its lowest peak on planet Quake?
Note 1: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Note 2: The values and planet names overlined in the question will change for each repetition in this part of the study. Please make sure you read the question carefully and familiarize yourself with the table headers or the graph framework before clicking the "Display data" button each time.
Proceed to graph framework
Timer: 11 seconds
[43/55]


Question: When was unemployment at its lowest peak on planet Quake?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[43/55]


Question: When was unemployment at its lowest peak on planet Quake?
Table: Evolution of Unemployment Rate on 4 Planets
February 2001
April 2002
June 2003
August 2004
October 2005
Timer: 10 seconds
[43/55]

x1


Question: When was unemployment at its lowest peak on planet Kreno?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[44/55]


Question: When was unemployment at its lowest peak on planet Kreno?
Evolution of Unemployment Rate on 4 Planets
July 2006
January 2007
July 2007
January 2008
July 2008
Timer: 10 seconds
[44/55]

x5

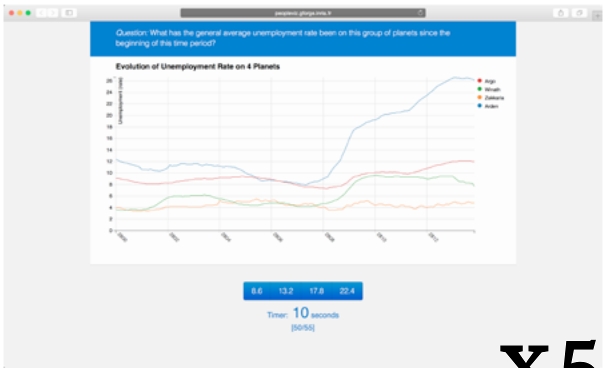

Chart type: line
Question: What has the general average unemployment rate been on this group of planets since the beginning of this time period?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[49/55]


Question: What has the general average unemployment rate been on this group of planets since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[49/55]

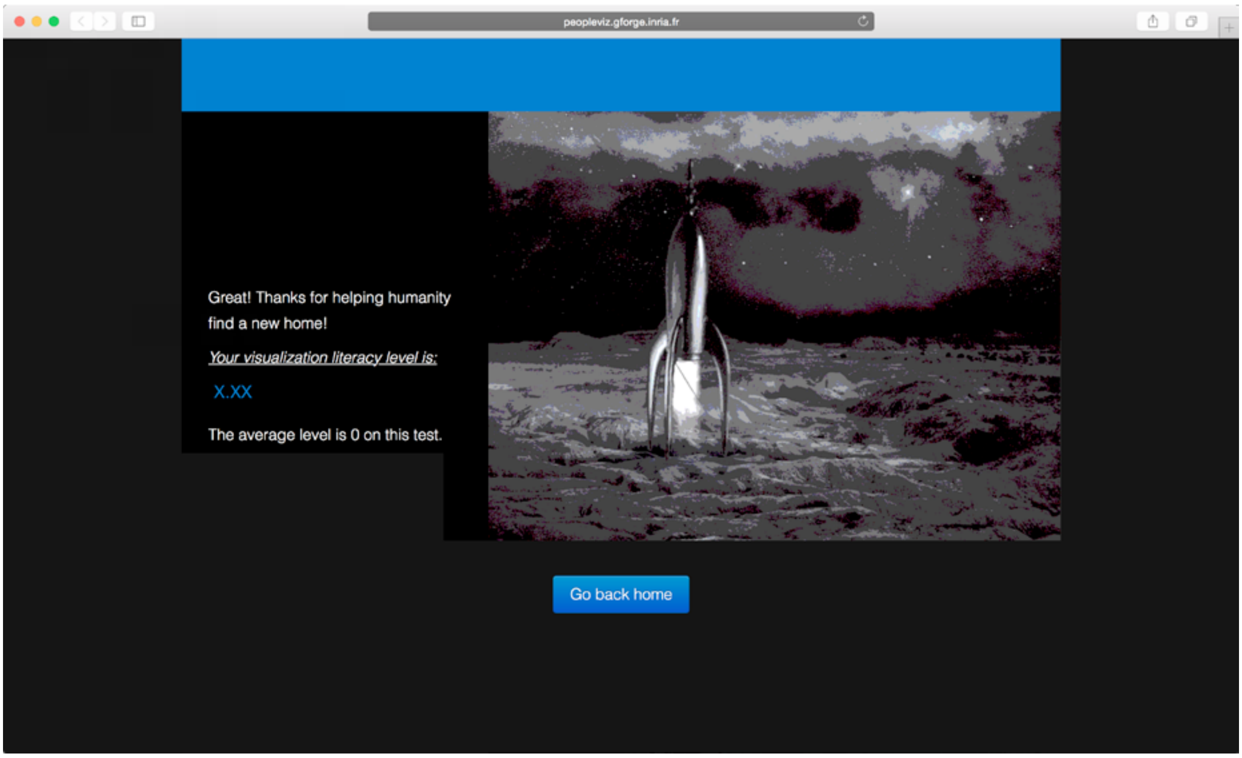

Question: What has the general average unemployment rate been on this group of planets since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
4.8
6.6
8.4
10.2
Timer: 10 seconds
[49/55]

x1


Question: What has the general average unemployment rate been on this group of planets since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[50/55]


Question: What has the general average unemployment rate been on this group of planets since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
8.6
13.2
17.8
22.4
Timer: 10 seconds
[50/55]

x5


peopleviz.gforge.inria.fr
Great! Thanks for helping humanity find a new home!
Your visualization literacy level is:
X.XX
The average level is 0 on this test.
Go back home

Appendix F

Visualization Literacy: Line Graphs Test 2

The full test is available at http://peopleviz.gforge.inria.fr/trunk/vLiteracy/home/tests/lg2/.

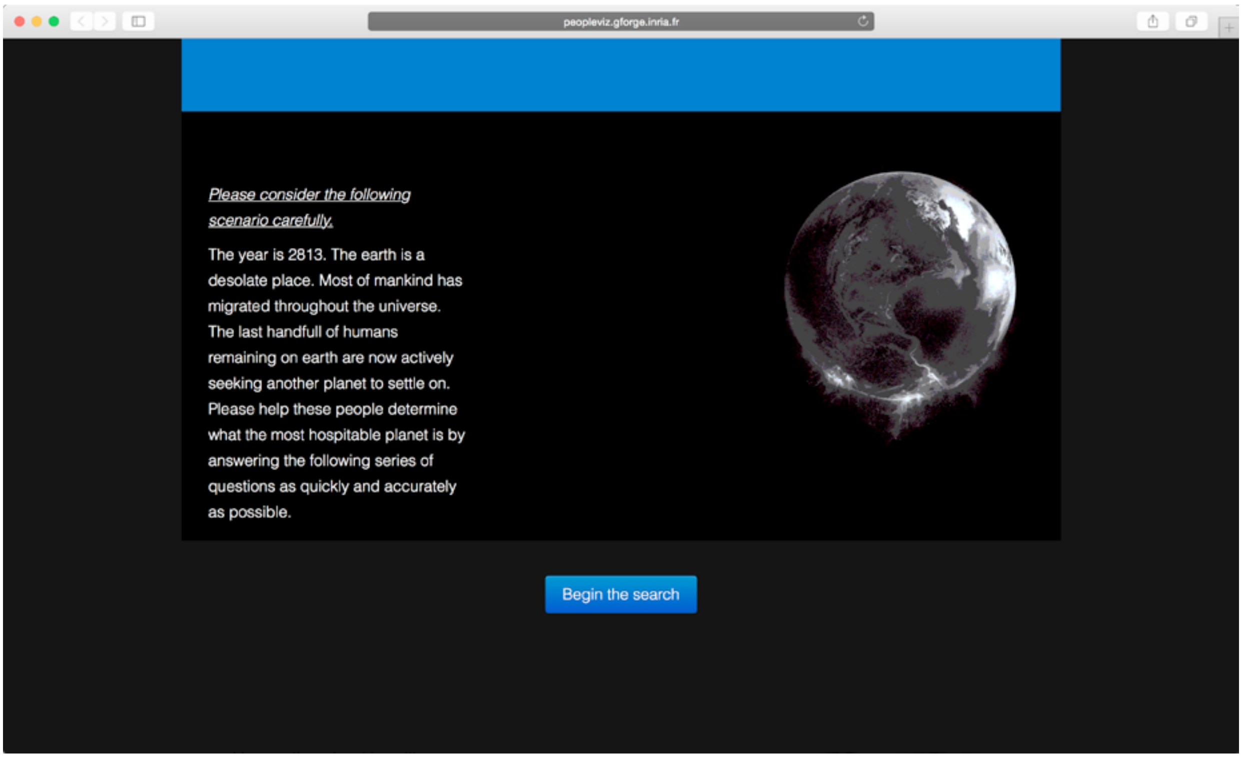
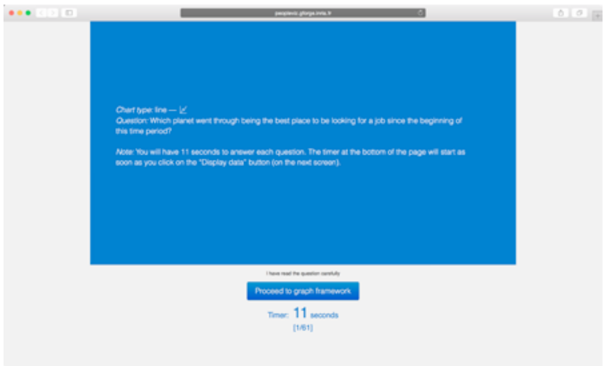


FIGURE F.1: The 10 different test-items in LG2.

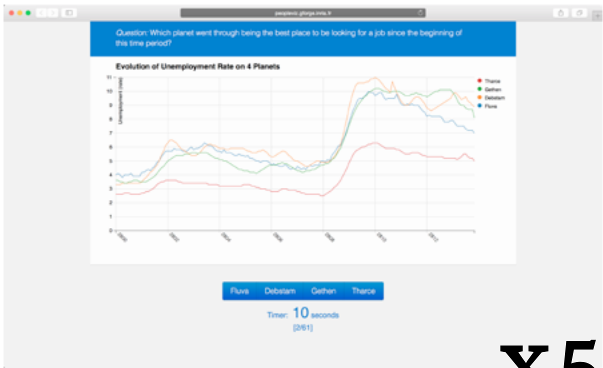

Question: Which planet went through being the best place to be looking for a job since the beginning of this time period?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[1/61]


Question: Which planet went through being the best place to be looking for a job since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[1/61]


Question: Which planet went through being the best place to be looking for a job since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Angor
Ungara
Vulcan
Lallor
Timer: 10 seconds
[1/61]

x1


Question: Which planet went through being the best place to be looking for a job since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[2/61]


Question: Which planet went through being the best place to be looking for a job since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Fluve
Debslam
Gethen
Tharce
Timer: 10 seconds
[2/61]

x5

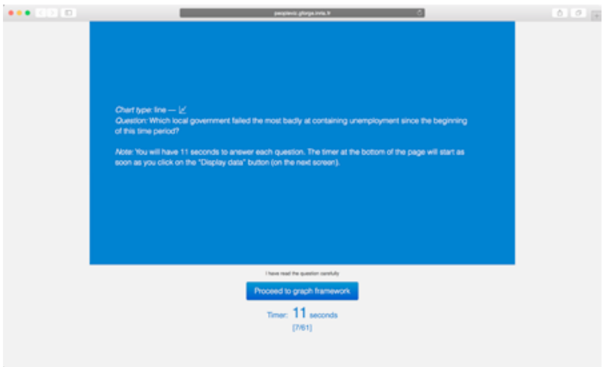


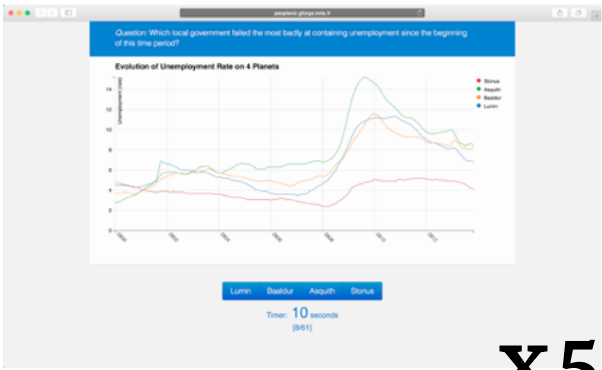
Chart type: line —
Question: Which local government failed the most badly at containing unemployment since the beginning of this time period?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[7/61]


Question: Which local government failed the most badly at containing unemployment since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[7/61]


Question: Which local government failed the most badly at containing unemployment since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Calafia
Dyofor
Talok
Margiot
Timer: 10 seconds
[7/61]

x 1


Question: Which local government failed the most badly at containing unemployment since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[8/61]


Question: Which local government failed the most badly at containing unemployment since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Lumin
Baaldur
Asquith
Stonus
Timer: 10 seconds
[8/61]

x 5


Chart type: line —
Question: Where did unemployment first go above 10.1?
Note 1: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Note 2: The values and planet names overlined in the question will change for each repetition in this part of the study. Please make sure you read the question carefully and familiarize yourself with the table headers or the graph framework before clicking the "Display data" button each time.
I have read the question carefully
Proceed to graph framework
Timer: 11 seconds
[13/61]


Question: Where did unemployment first go above 10.1?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[13/61]


Question: Where did unemployment first go above 10.1?
Table: Evolution of Unemployment Rate on 4 Planets
Strontia
Maeg
Salinn
Deneb
Timer: 10 seconds
[13/61]

x1


Question: Where did unemployment first go above 18.8?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[14/61]


Question: Where did unemployment first go above 18.8?
Evolution of Unemployment Rate on 4 Planets
Zamaron
Gallea
Helicon
Orsha
Timer: 10 seconds
[14/61]

x5

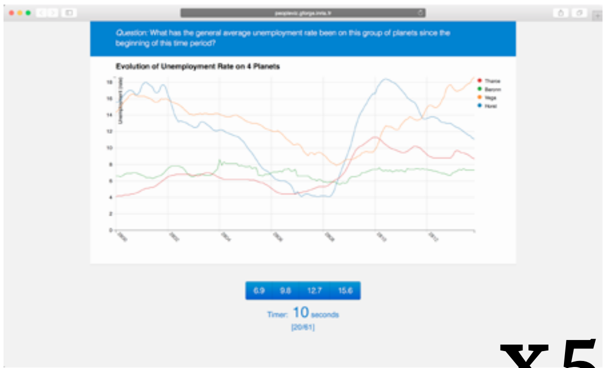

Question: What has the general average unemployment rate been on this group of planets since the beginning of this time period?
Proceed to graph framework
Timer: 11 seconds


Question: What has the general average unemployment rate been on this group of planets since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds


Question: What has the general average unemployment rate been on this group of planets since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
4.8
6.6
8.4
10.2
Timer: 10 seconds

x1


Question: What has the general average unemployment rate been on this group of planets since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds


Question: What has the general average unemployment rate been on this group of planets since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
6.9
9.8
12.7
15.6
Timer: 10 seconds

x5

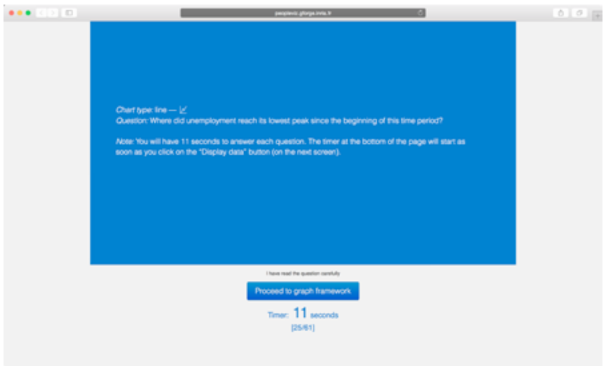


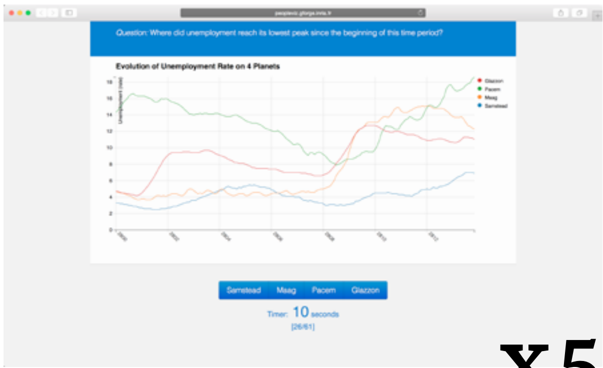
Chart type: line
Question: Where did unemployment reach its lowest peak since the beginning of this time period?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[25/61]


Question: Where did unemployment reach its lowest peak since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[25/61]


Question: Where did unemployment reach its lowest peak since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Garon
Brasl
Dysia
Darlbow
Timer: 10 seconds
[25/61]

x1


Question: Where did unemployment reach its lowest peak since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[26/61]


Question: Where did unemployment reach its lowest peak since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Samstead
Maeg
Pacem
Glazzon
Timer: 10 seconds
[26/61]

x5

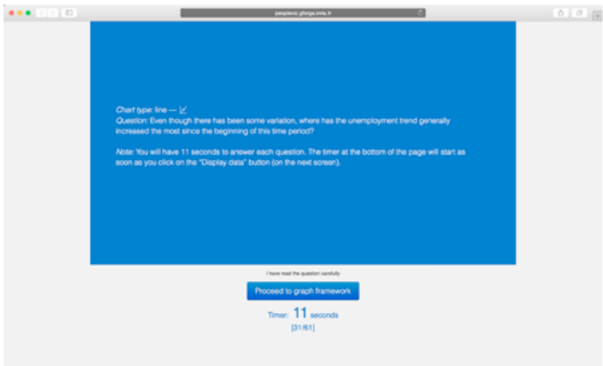
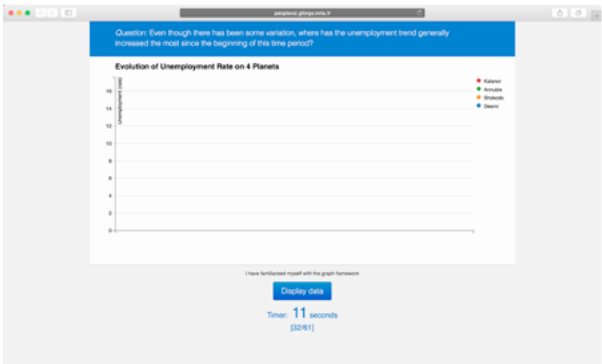


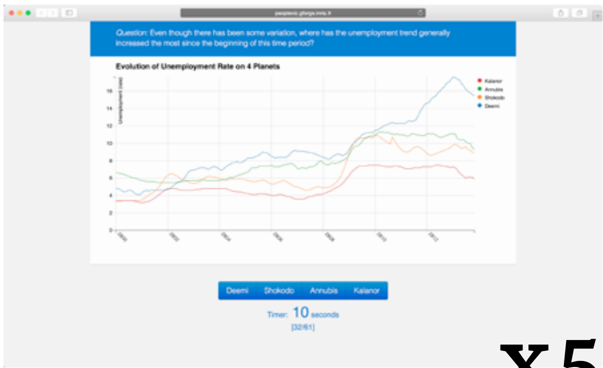
Chart type: line
Question: Even though there has been some variation, where has the unemployment trend generally increased the most since the beginning of this time period?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[31/61]


Question: Even though there has been some variation, where has the unemployment trend generally increased the most since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[31/61]


Question: Even though there has been some variation, where has the unemployment trend generally increased the most since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Krypton
Comagrave
Brasl
Arcadia
Timer: 10 seconds
[31/61]

x1


Question: Even though there has been some variation, where has the unemployment trend generally increased the most since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[32/61]


Question: Even though there has been some variation, where has the unemployment trend generally increased the most since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Deemi
Shokodo
Annubis
Kalanor
Timer: 10 seconds
[32/61]

x5

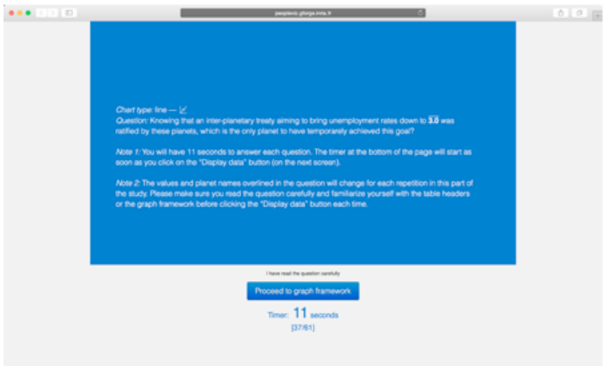


Chart type: line
Question: Knowing that an inter-planetary treaty aiming to bring unemployment rates down to 3.0 was ratified by these planets, which is the only planet to have temporarily achieved this goal?
Proceed to graph framework
Timer: 11 seconds


Question: Knowing that an inter-planetary treaty aiming to bring unemployment rates down to 3.0 was ratified by these planets, which is the only planet to have temporarily achieved this goal?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds


Question: Knowing that an inter-planetary treaty aiming to bring unemployment rates down to 3.0 was ratified by these planets, which is the only planet to have temporarily achieved this goal?
Table: Evolution of Unemployment Rate on 4 Planets
Amerac
Strontia
Pandora
Rann
Timer: 10 seconds

x1


Question: Knowing that an inter-planetary treaty aiming to bring unemployment rates down to 2.9 was ratified by these planets, which is the only planet to have temporarily achieved this goal?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds


Question: Knowing that an inter-planetary treaty aiming to bring unemployment rates down to 2.9 was ratified by these planets, which is the only planet to have temporarily achieved this goal?
Evolution of Unemployment Rate on 4 Planets
Tanith
Darlbow
Ria
Cil
Timer: 10 seconds

x5

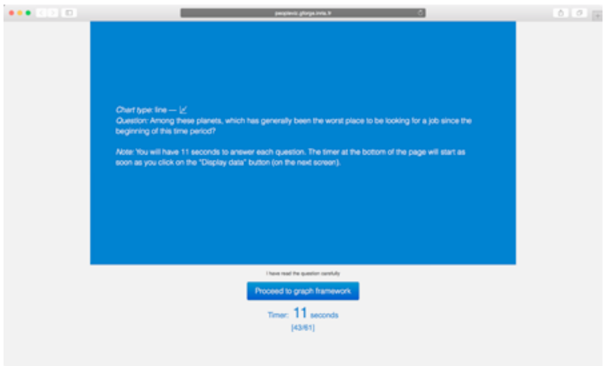
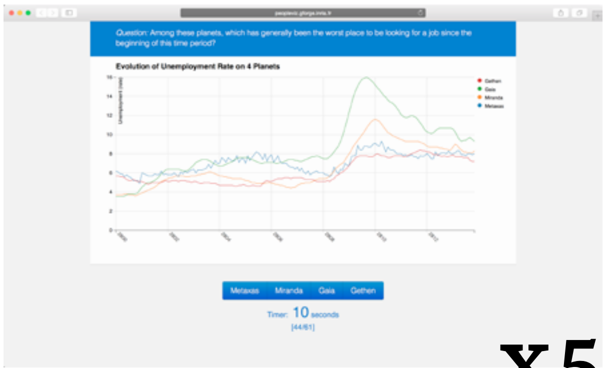

Proceed to graph framework
Timer: 11 seconds
[43/61]


Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[43/61]


Table: Evolution of Unemployment Rate on 4 Planets
Rokanan
Zoranel
Naltor
Ciress
Timer: 10 seconds
[43/61]

x1


Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[44/61]


Evolution of Unemployment Rate on 4 Planets
Metaxas
Miranda
Gaia
Gethen
Timer: 10 seconds
[44/61]

x5

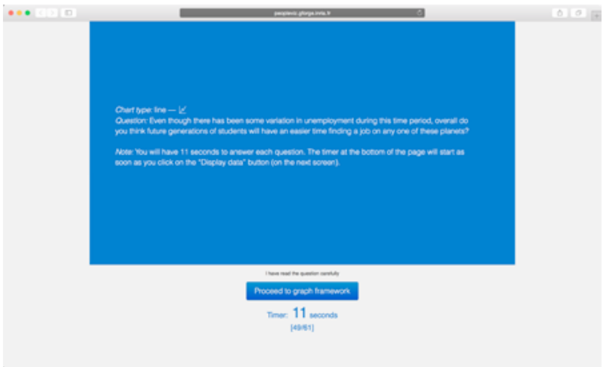


Chart type: line —
Question: Even though there has been some variation in unemployment during this time period, overall do you think future generations of students will have an easier time finding a job on any one of these planets?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[49/61]


Question: Even though there has been some variation in unemployment during this time period, overall do you think future generations of students will have an easier time finding a job on any one of these planets?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[49/61]


Question: Even though there has been some variation in unemployment during this time period, overall do you think future generations of students will have an easier time finding a job on any one of these planets?
Table: Evolution of Unemployment Rate on 4 Planets
Yes
No
Timer: 10 seconds
[49/61]

x1


Question: Even though there has been some variation in unemployment during this time period, overall do you think future generations of students will have an easier time finding a job on any one of these planets?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[50/61]


Question: Even though there has been some variation in unemployment during this time period, overall do you think future generations of students will have an easier time finding a job on any one of these planets?
Evolution of Unemployment Rate on 4 Planets
Yes
No
Timer: 10 seconds
[50/61]

x5

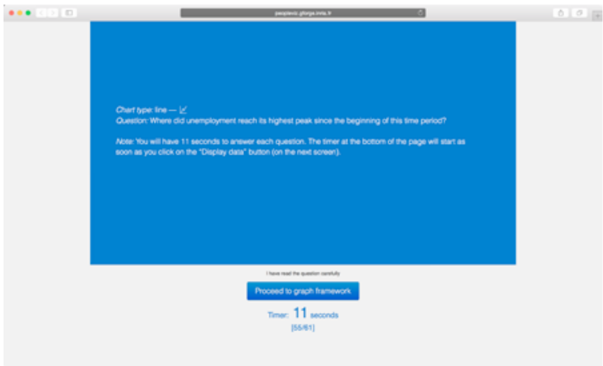
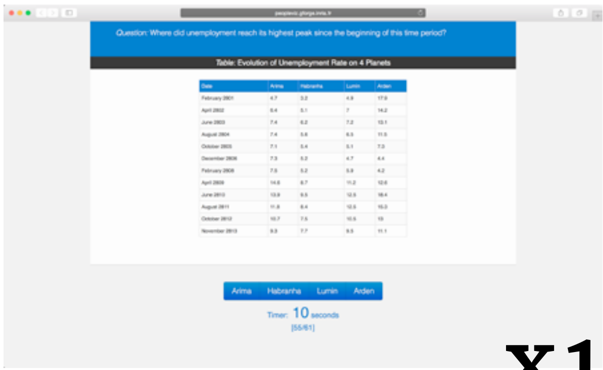


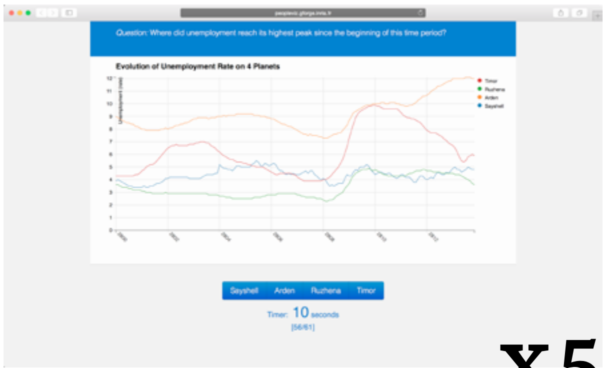
Chart type: line —
Question: Where did unemployment reach its highest peak since the beginning of this time period?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[55/61]


Question: Where did unemployment reach its highest peak since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[55/61]

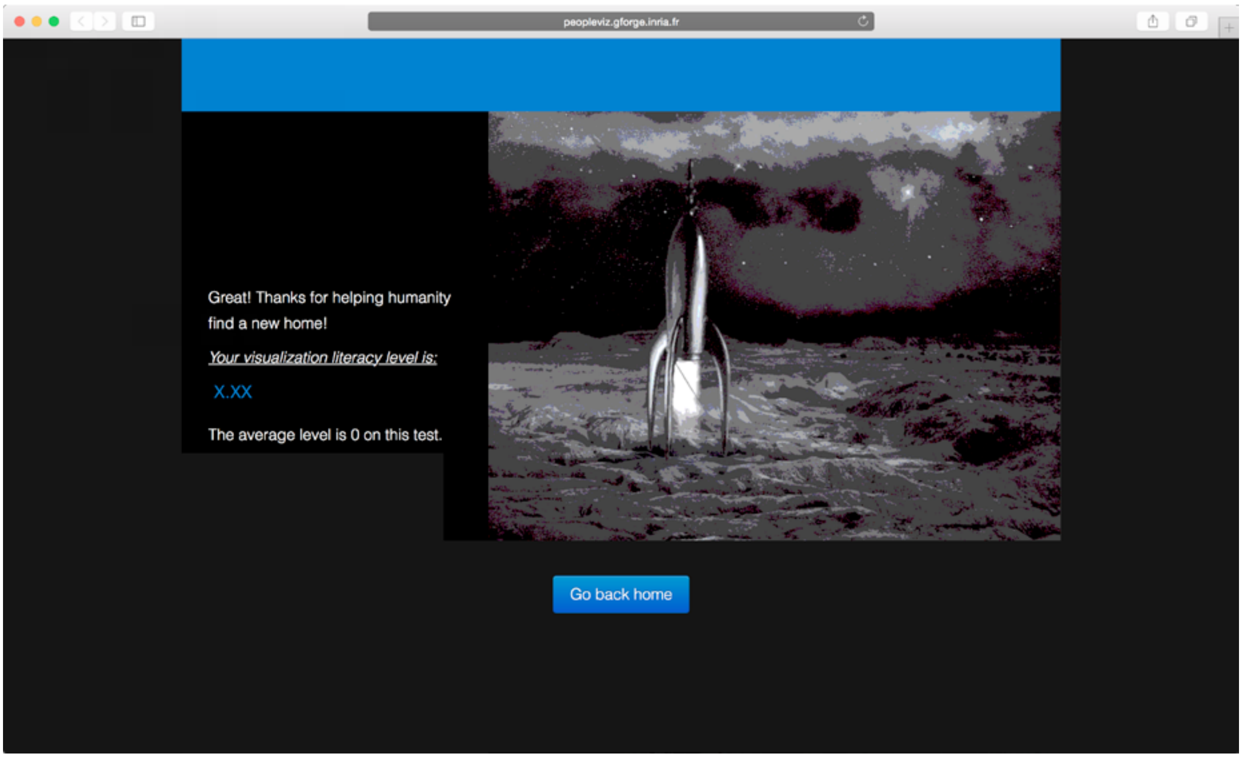

Question: Where did unemployment reach its highest peak since the beginning of this time period?
Table: Evolution of Unemployment Rate on 4 Planets
Arima
Habranha
Lumin
Arden
Timer: 10 seconds
[55/61]

x1


Question: Where did unemployment reach its highest peak since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Display data
Timer: 11 seconds
[56/61]


Question: Where did unemployment reach its highest peak since the beginning of this time period?
Evolution of Unemployment Rate on 4 Planets
Seyshell
Arden
Ruzhena
Timor
Timer: 10 seconds
[56/61]

x5


peopleviz.gforge.inria.fr
Great! Thanks for helping humanity find a new home!
Your visualization literacy level is:
X.XX
The average level is 0 on this test.
Go back home

Appendix G

Visualization Literacy: Bar Charts Test

The full test is available at http://peopleviz.gforge.inria.fr/trunk/vLiteracy/home/tests/bc/.

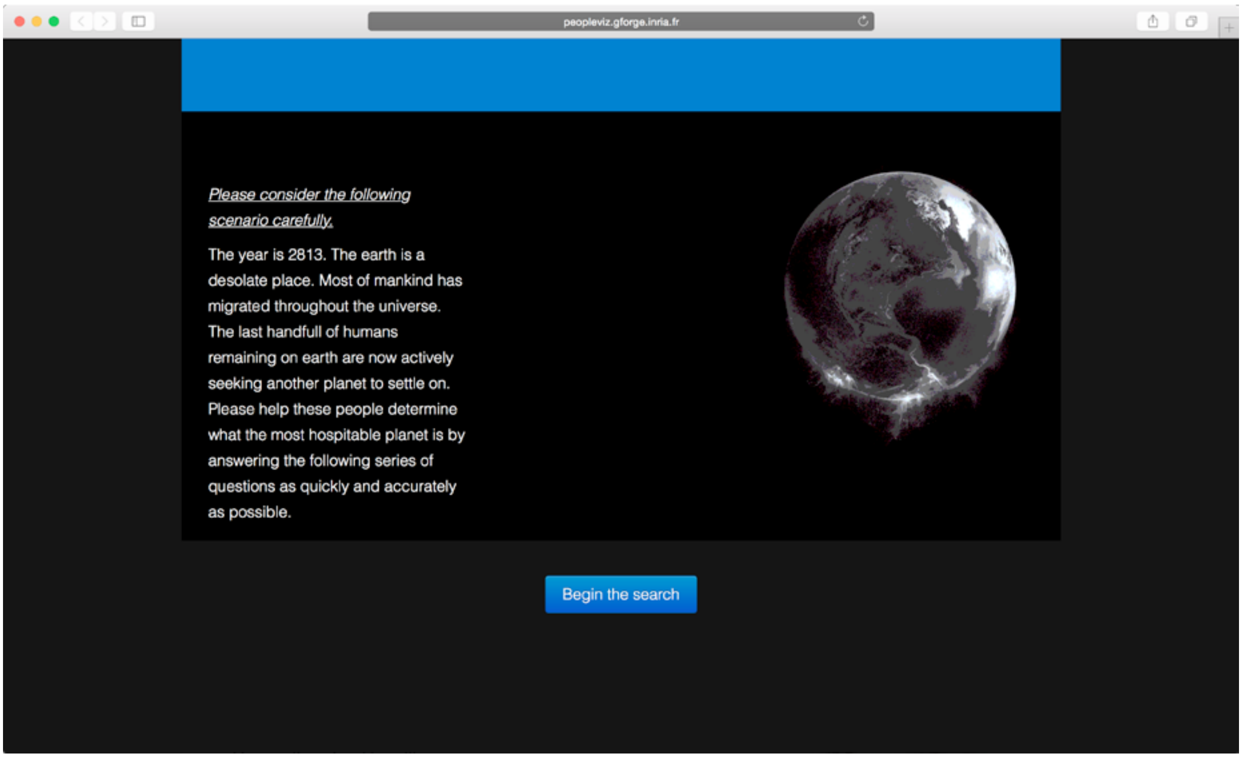


FIGURE G.1: The 8 different test-items in BC.

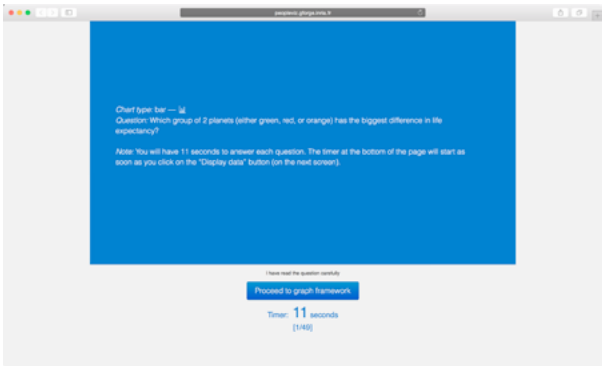
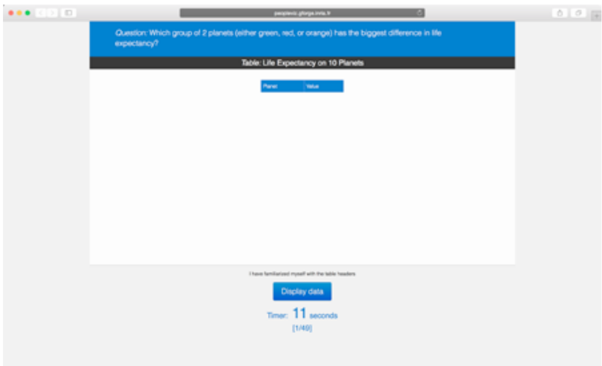
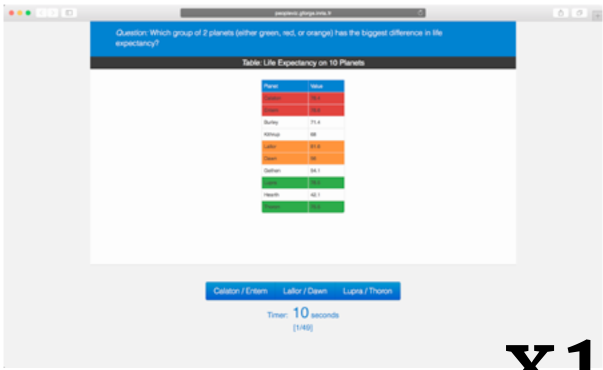


Chart type: bar
Question: Which group of 2 planets (either green, red, or orange) has the biggest difference in life expectancy?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[1/49]


Question: Which group of 2 planets (either green, red, or orange) has the biggest difference in life expectancy?
Table: Life Expectancy on 10 Planets
Display data
Timer: 11 seconds
[1/49]

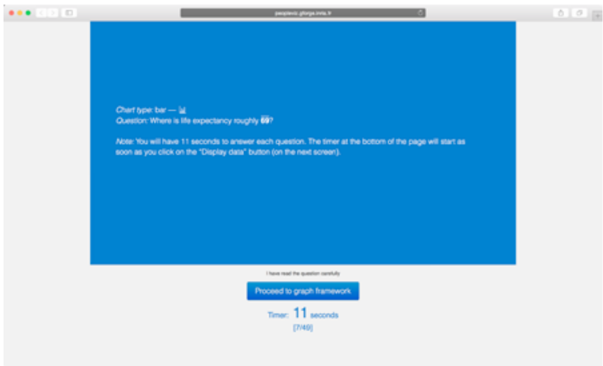

Question: Which group of 2 planets (either green, red, or orange) has the biggest difference in life expectancy?
Table: Life Expectancy on 10 Planets
Calaton / Entem
Lallor / Dawn
Lupra / Thoron
Timer: 10 seconds
[1/49]

x1

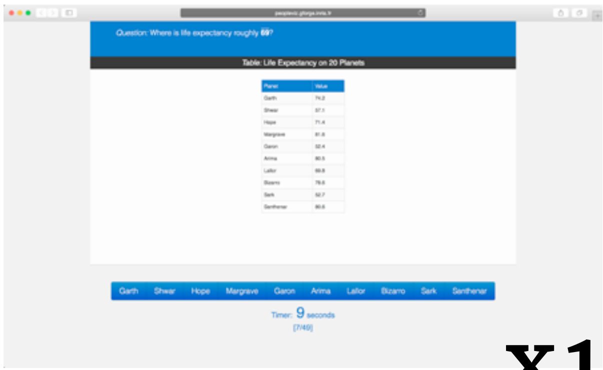

Question: Which group of 2 planets (either green, red, or orange) has the biggest difference in life expectancy?
Life Expectancy on 10 Planets
Display data
Timer: 11 seconds
[2/49]


Question: Which group of 2 planets (either green, red, or orange) has the biggest difference in life expectancy?
Life Expectancy on 10 Planets
Nator / Sarkovy
Livia / Rys
Miranda / Synnax
Timer: 10 seconds
[2/49]

x5

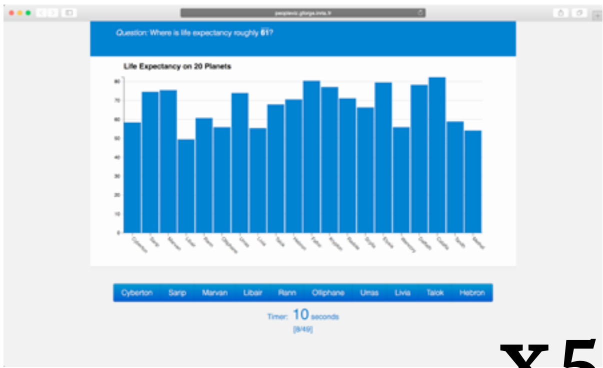

Question: Where is life expectancy roughly 69?
Proceed to graph framework
Timer: 11 seconds
[7/49]


Question: Where is life expectancy roughly 69?
Table: Life Expectancy on 20 Planets
Display data
Timer: 11 seconds
[7/49]

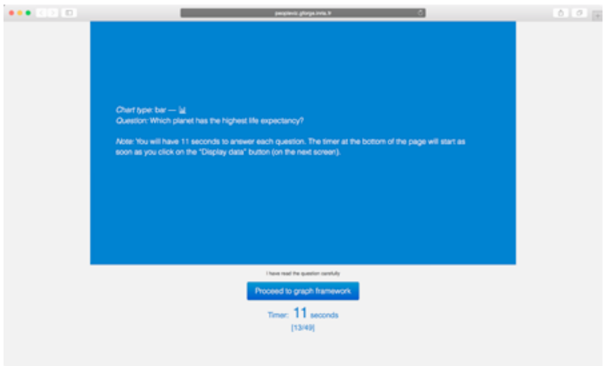

Question: Where is life expectancy roughly 69?
Table: Life Expectancy on 20 Planets
Garth
Shwar
Hope
Margrave
Garon
Arima
Lallor
Bizarro
Sark
Santhenar
Timer: 9 seconds
[7/49]

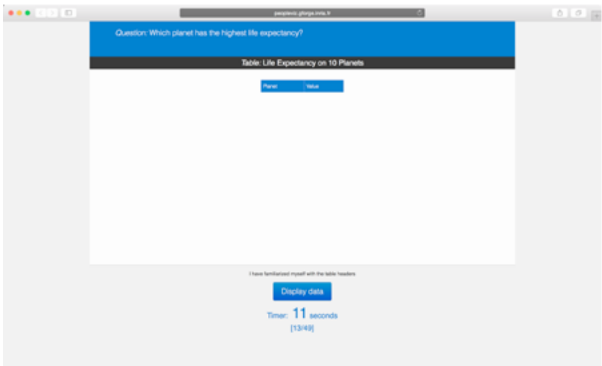
x1

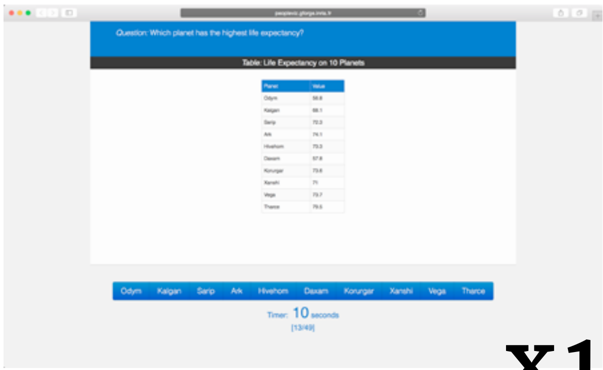

Question: Where is life expectancy roughly 61?
Life Expectancy on 20 Planets
Display data
Timer: 11 seconds
[8/49]


Question: Where is life expectancy roughly 61?
Life Expectancy on 20 Planets
Cyberton
Sarp
Marvan
Libair
Rann
Oliphane
Urras
Livia
Talok
Hebron
Timer: 10 seconds
[8/49]

x5

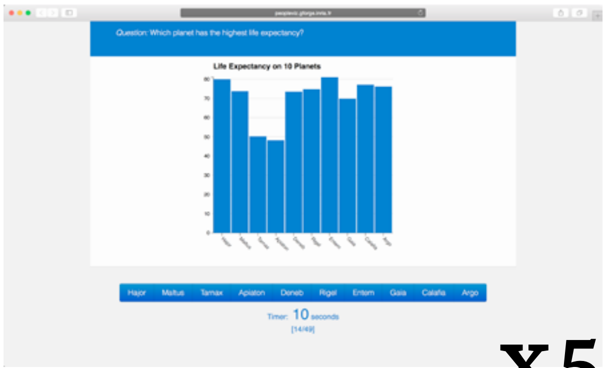

Question: Which planet has the highest life expectancy?
Proceed to graph framework
Timer: 11 seconds
[13/49]


Question: Which planet has the highest life expectancy?
Table: Life Expectancy on 10 Planets
Display data
Timer: 11 seconds
[13/49]


Question: Which planet has the highest life expectancy?
Table: Life Expectancy on 10 Planets
Odym
Kalgan
Sarip
Ark
Hivehom
Daxam
Korurgar
Xanshi
Vega
Tharce
Timer: 10 seconds
[13/49]

x1


Question: Which planet has the highest life expectancy?
Life Expectancy on 10 Planets
Display data
Timer: 11 seconds
[14/49]


Question: Which planet has the highest life expectancy?
Life Expectancy on 10 Planets
Hajor
Maltus
Tamax
Aplaton
Deneb
Rigel
Entom
Gaia
Calafia
Argo
Timer: 10 seconds
[14/49]

x5

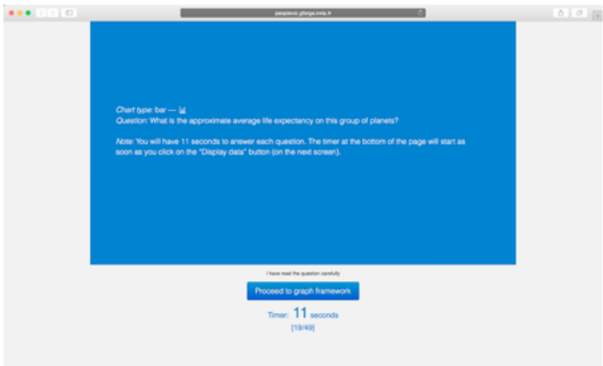


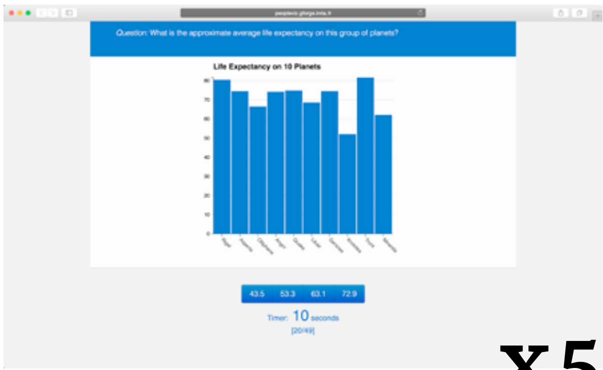
Chart type: bar —
Question: What is the approximate average life expectancy on this group of planets?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
I have read the question carefully
Proceed to graph framework
Timer: 11 seconds
[19/49]


Question: What is the approximate average life expectancy on this group of planets?
Table: Life Expectancy on 10 Planets
Planet
Value
Display data
Timer: 11 seconds
[19/49]

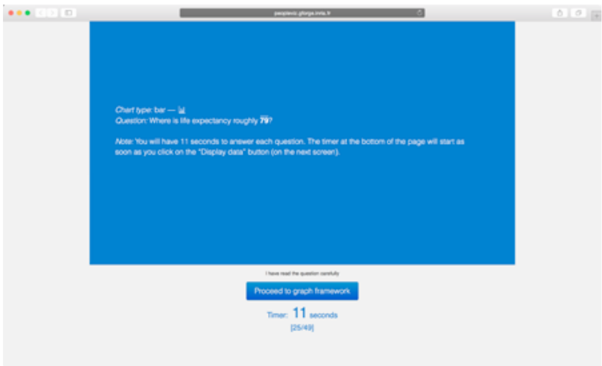

Question: What is the approximate average life expectancy on this group of planets?
Table: Life Expectancy on 10 Planets
41.1
50.6
60.1
69.9
Timer: 10 seconds
[19/49]

x1

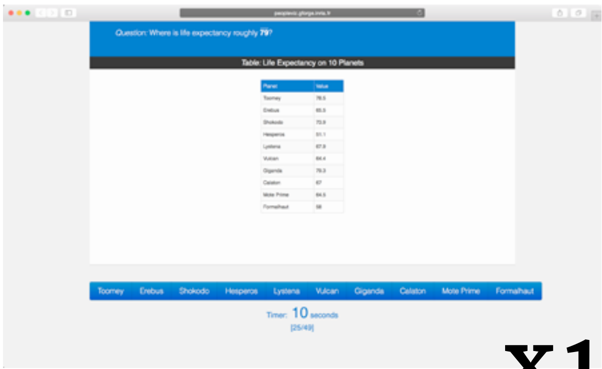

Question: What is the approximate average life expectancy on this group of planets?
Life Expectancy on 10 Planets
Display data
Timer: 11 seconds
[20/49]


Question: What is the approximate average life expectancy on this group of planets?
Life Expectancy on 10 Planets
43.5
53.3
63.1
72.9
Timer: 10 seconds
[20/49]

x5

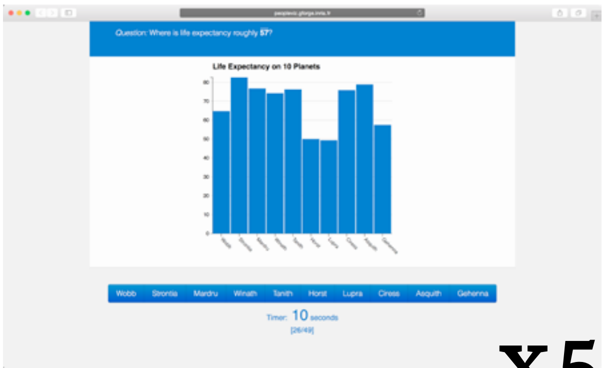

Proceed to graph framework
Timer: 11 seconds
[25/49]


Question: Where is life expectancy roughly 79?
Table: Life Expectancy on 10 Planets
Display data
Timer: 11 seconds
[25/49]

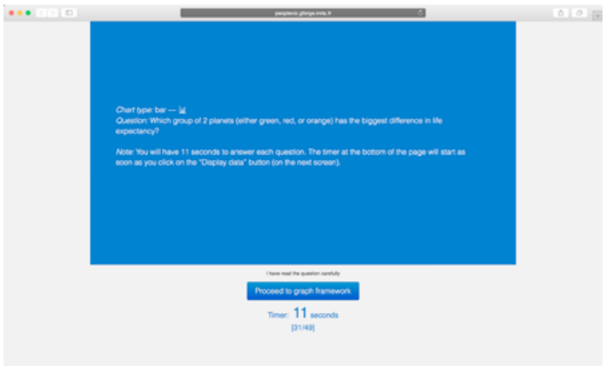

Question: Where is life expectancy roughly 79?
Table: Life Expectancy on 10 Planets
Toomey
Erebus
Shokodo
Hesperos
Lystena
Vulcan
Giganda
Calaton
Mote Prime
Formalhaut
Timer: 10 seconds
[25/49]

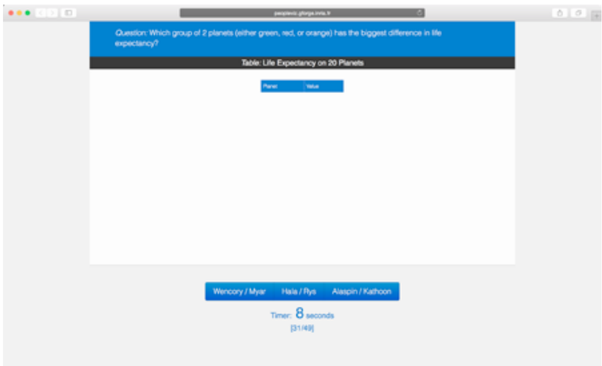
x1

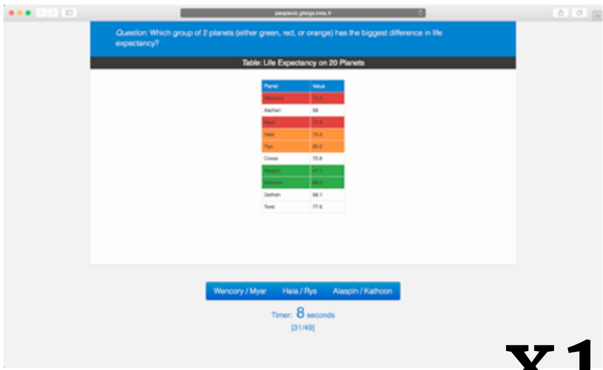

Life Expectancy on 10 Planets
Display data
Timer: 11 seconds
[26/49]


Life Expectancy on 10 Planets
Timer: 10 seconds
[26/49]

x5







x1

x5

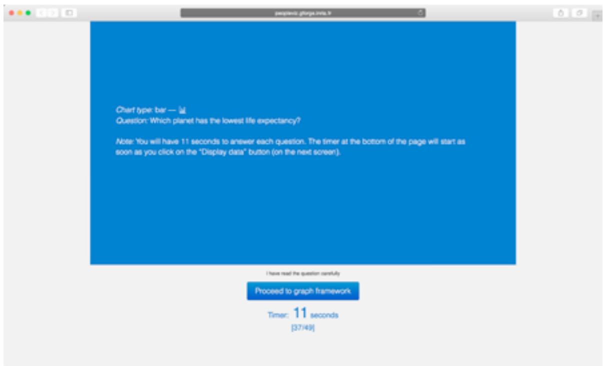
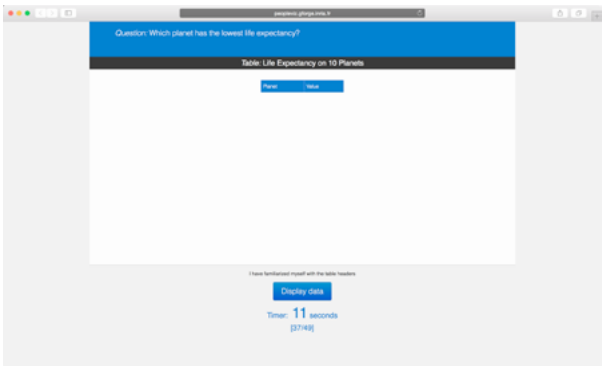
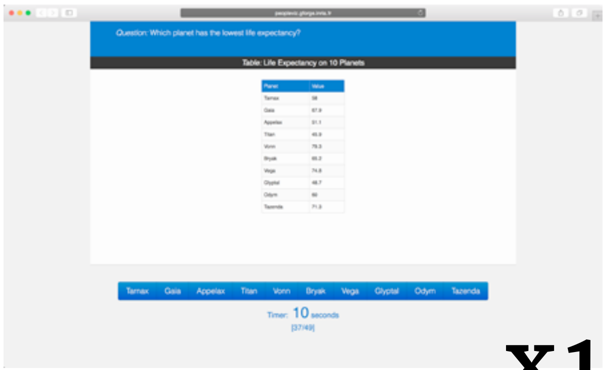


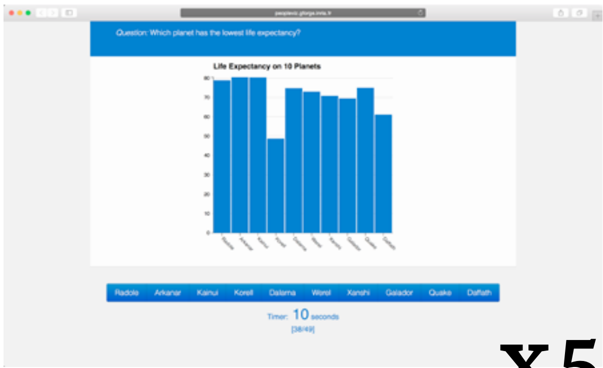
Chart type: bar
Question: Which planet has the lowest life expectancy?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[37/49]


Question: Which planet has the lowest life expectancy?
Table: Life Expectancy on 10 Planets
Display data
Timer: 11 seconds
[37/49]

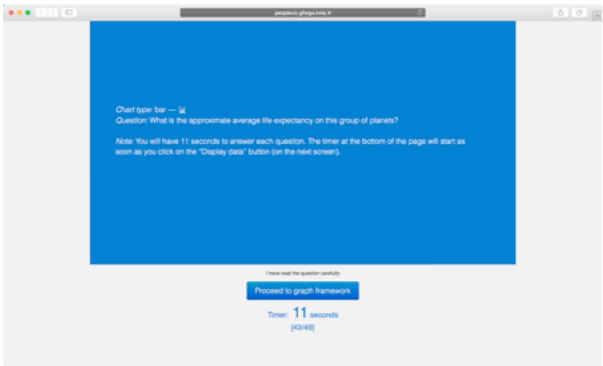

Question: Which planet has the lowest life expectancy?
Table: Life Expectancy on 10 Planets
Tarnax Gaia Appelax Titan Vonn Bryak Vega Glyptal Odym Tazenda
Timer: 10 seconds
[37/49]

x1

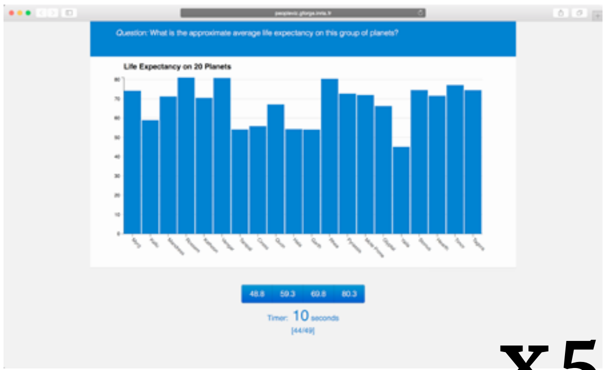

Question: Which planet has the lowest life expectancy?
Life Expectancy on 10 Planets
Display data
Timer: 11 seconds
[38/49]


Question: Which planet has the lowest life expectancy?
Life Expectancy on 10 Planets
Radole Arkanar Kainui Koreil Dalama Wersl Xanshi Galador Quake Daftath
Timer: 10 seconds
[38/49]

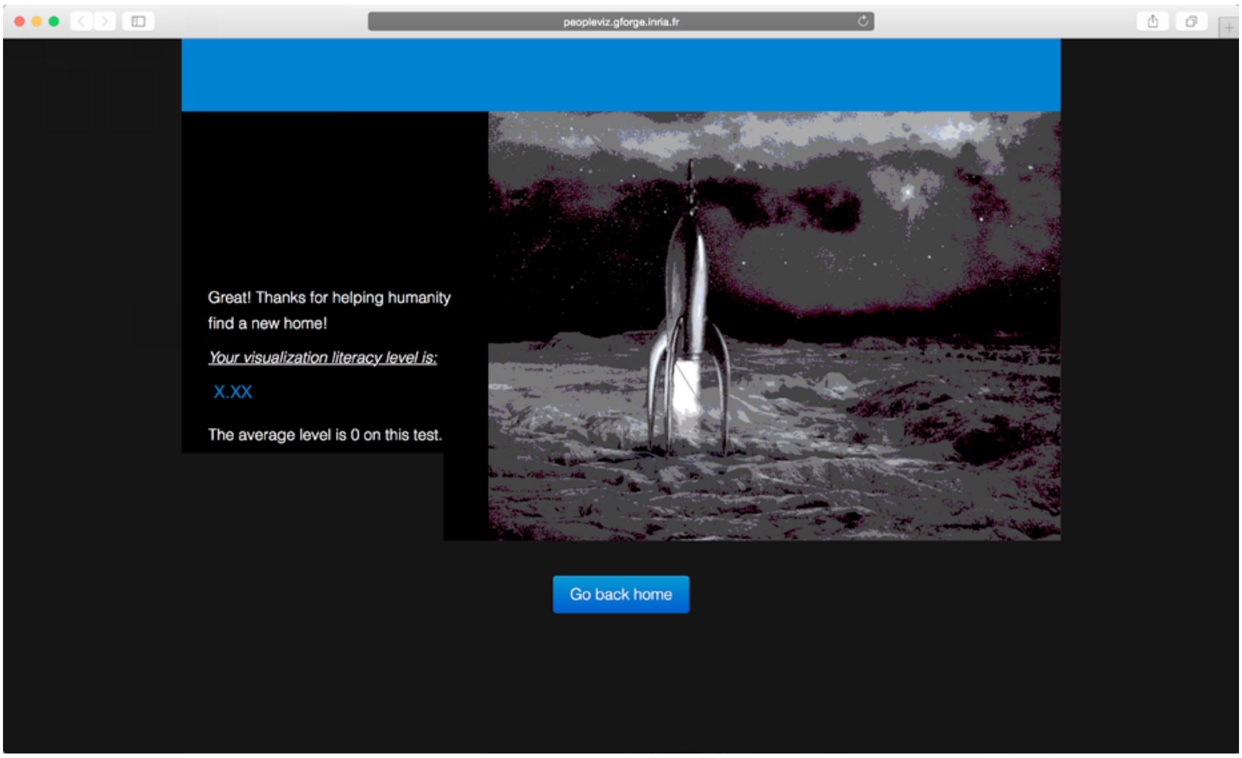
x5



x1



x5


peopleviz.gforge.inria.fr
Great! Thanks for helping humanity find a new home!
Your visualization literacy level is:
X.XX
The average level is 0 on this test.
Go back home

Appendix H

Visualization Literacy: Scatterplots Test

The full test is available at http://peopleviz.gforge.inria.fr/trunk/vLiteracy/home/tests/sp/.

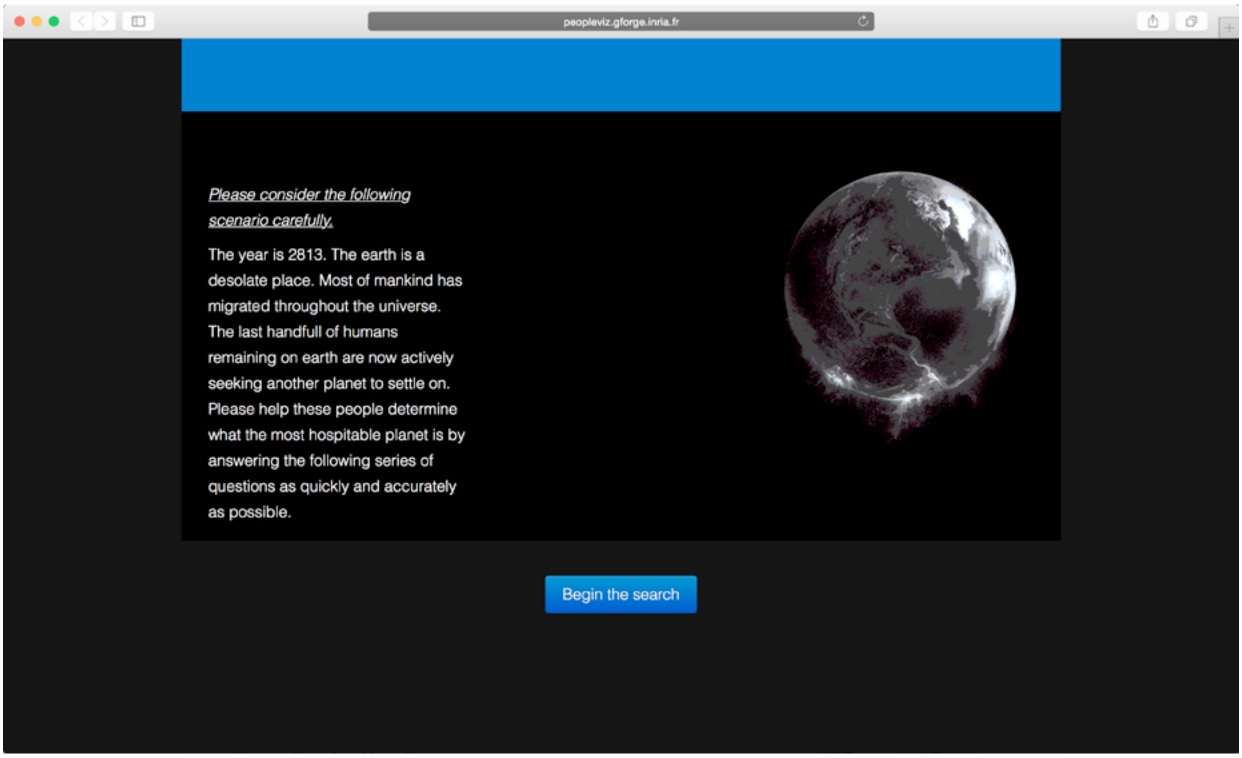


FIGURE H.1: The 10 different test-items in SP.

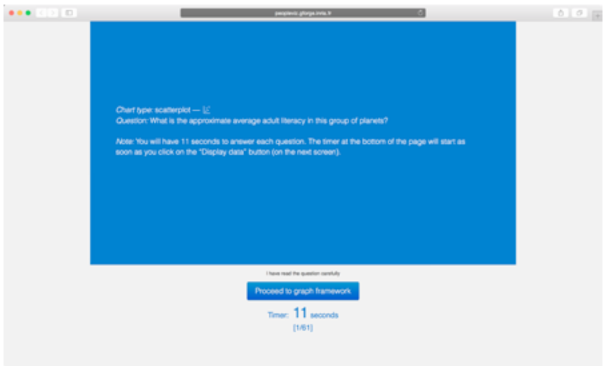


Chart type: scatterplot
Question: What is the approximate average adult literacy in this group of planets?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[1/61]


Question: What is the approximate average adult literacy in this group of planets?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds
[1/61]

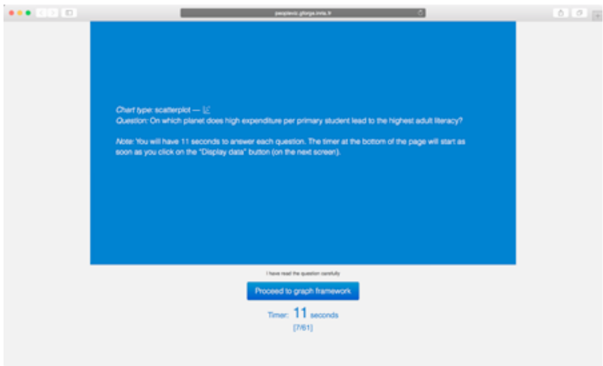

Question: What is the approximate average adult literacy in this group of planets?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
38.1
53
67.9
82.8
Timer: 10 seconds
[1/61]

x1


Question: What is the approximate average adult literacy in this group of planets?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds
[2/61]


Question: What is the approximate average adult literacy in this group of planets?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
66.8
74.7
82.6
90.5
Timer: 10 seconds
[2/61]

x5

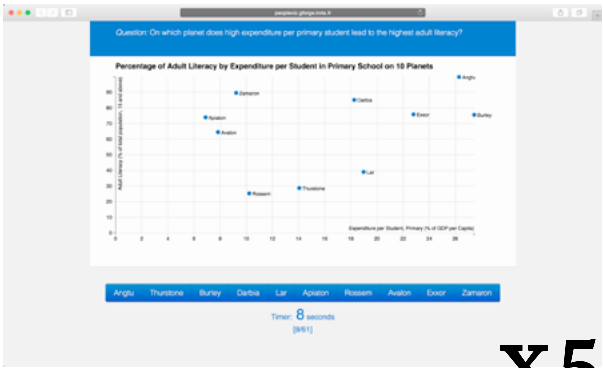

Question: On which planet does high expenditure per primary student lead to the highest adult literacy?
Proceed to graph framework
Timer: 11 seconds
[7/61]


Question: On which planet does high expenditure per primary student lead to the highest adult literacy?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds
[7/61]

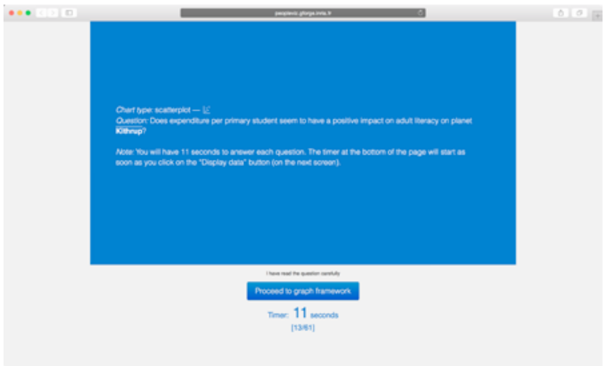

Question: On which planet does high expenditure per primary student lead to the highest adult literacy?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Brasl
Hiberlac
Kallu
Sarkovy
Kalanor
Cuthbert
Mardru
Daribow
Umas
Livia
Timer: 10 seconds
[7/61]

x1


Question: On which planet does high expenditure per primary student lead to the highest adult literacy?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds
[8/61]


Question: On which planet does high expenditure per primary student lead to the highest adult literacy?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Angtu
Thurstone
Burley
Darbia
Lar
Apiaton
Rossem
Avalon
Exxor
Zamaron
Timer: 8 seconds
[8/61]

x5


Proceed to graph framework
Timer: 11 seconds


Display data
Timer: 11 seconds


Yes No
Timer: 10 seconds

x1


Display data
Timer: 11 seconds


Yes No
Timer: 10 seconds

x5

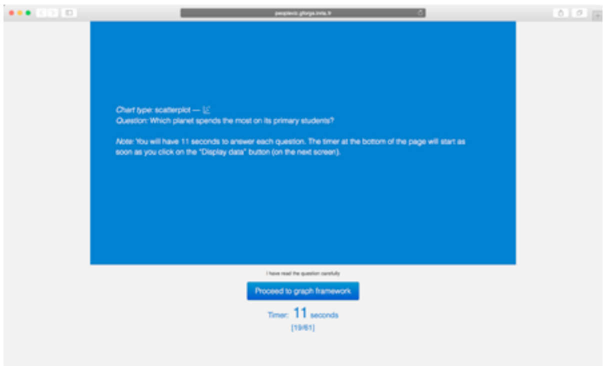


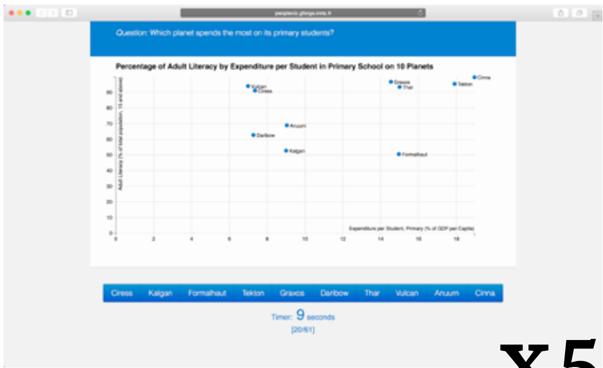
Chart type: scatterplot
Question: Which planet spends the most on its primary students?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[19/61]


Question: Which planet spends the most on its primary students?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds
[19/61]

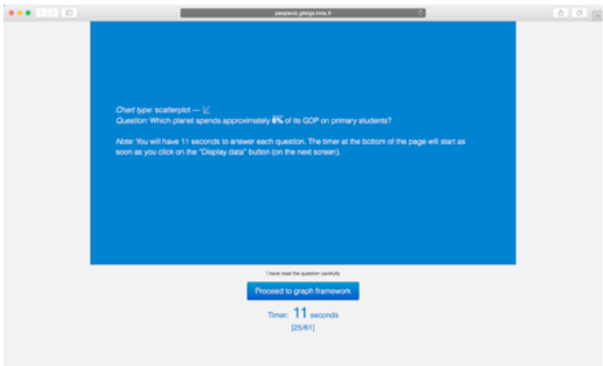

Question: Which planet spends the most on its primary students?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Garrota
Daxam
Lernista
Tun
Mendel
Zebron
Snatanni
Orando
Arkanar
Urras
Timer: 10 seconds
[19/61]

x1

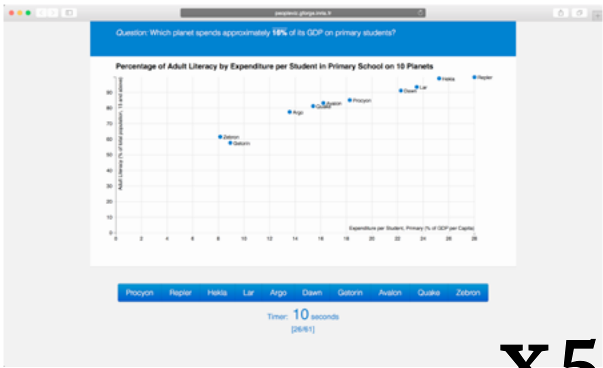

Question: Which planet spends the most on its primary students?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds
[20/61]


Question: Which planet spends the most on its primary students?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Ciress
Kalgan
Formalhaut
Tekton
Graxos
Darlbow
Thar
Vulcan
Anuum
Cirrna
Timer: 9 seconds
[20/61]

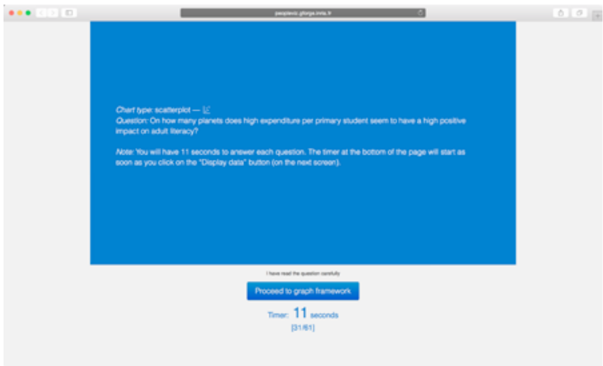
x5



x1



x5

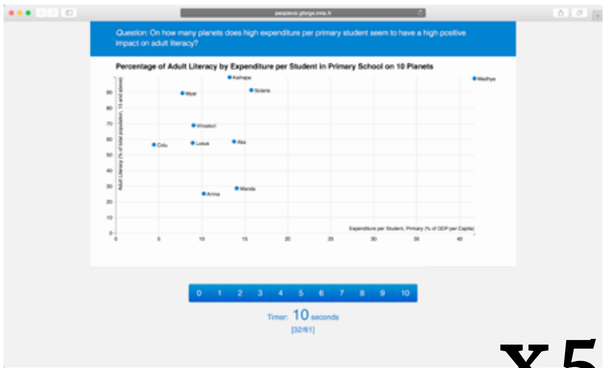

Proceed to graph framework
Timer: 11 seconds


Display data
Timer: 11 seconds

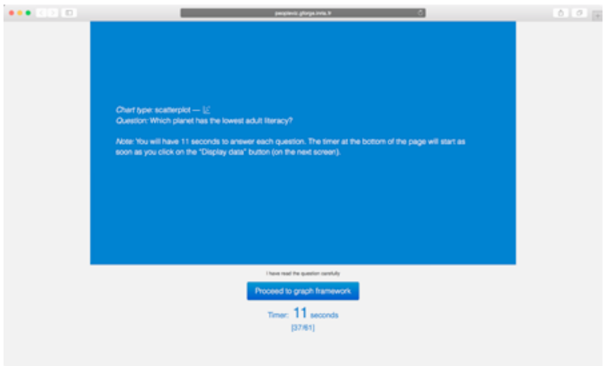

0 1 2 3 4 5 6 7 8 9 10
Timer: 10 seconds

x1


Display data
Timer: 11 seconds


0 1 2 3 4 5 6 7 8 9 10
Timer: 10 seconds

x5

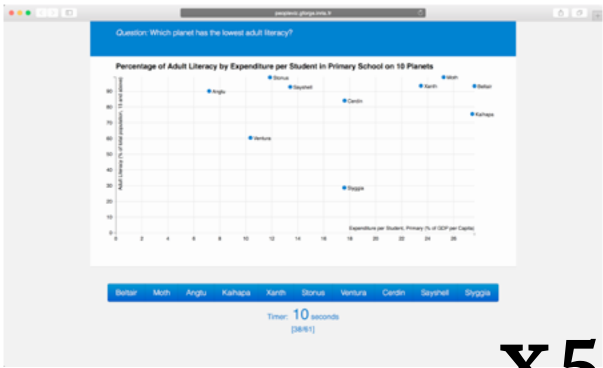

Question: Which planet has the lowest adult literacy?
Proceed to graph framework
Timer: 11 seconds


Question: Which planet has the lowest adult literacy?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds

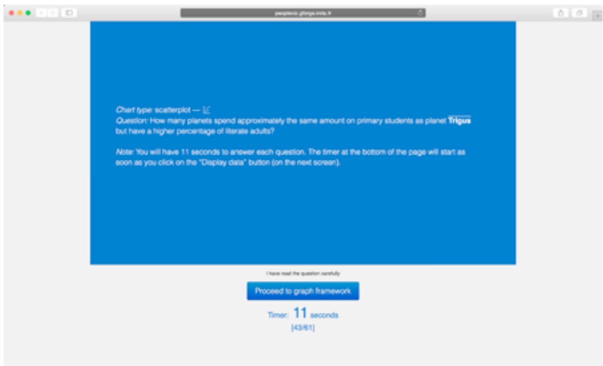

Question: Which planet has the lowest adult literacy?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Timer: 10 seconds

x1

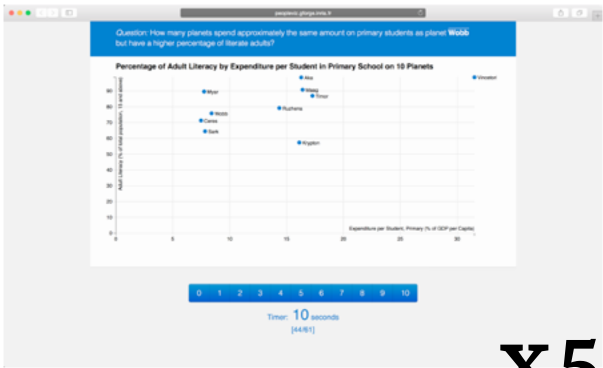

Question: Which planet has the lowest adult literacy?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds


Question: Which planet has the lowest adult literacy?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Timer: 10 seconds

x5



x1



x5

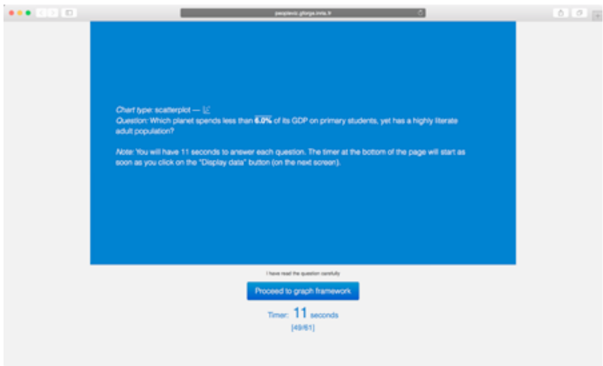


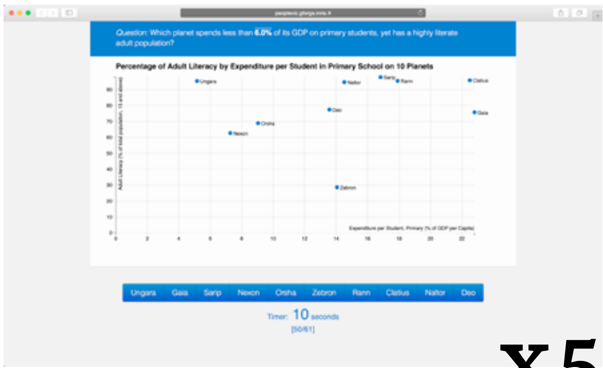
Chart type: scatterplot —
Question: Which planet spends less than 6.0% of its GDP on primary students, yet has a highly literate adult population?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[49/61]


Question: Which planet spends less than 6.0% of its GDP on primary students, yet has a highly literate adult population?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds
[49/61]


Question: Which planet spends less than 6.0% of its GDP on primary students, yet has a highly literate adult population?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Bolovax
Mogo
Contraria
Sheathclaws
Moth
Mote Prime
Werel
Shwar
Rys
Thordia
Timer: 10 seconds
[49/61]

x1


Question: Which planet spends less than 6.0% of its GDP on primary students, yet has a highly literate adult population?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds
[50/61]


Question: Which planet spends less than 6.0% of its GDP on primary students, yet has a highly literate adult population?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Ungara
Gaia
Sarp
Nexon
Orsha
Zebron
Rann
Clatus
Naltor
Deo
Timer: 10 seconds
[50/61]

x5

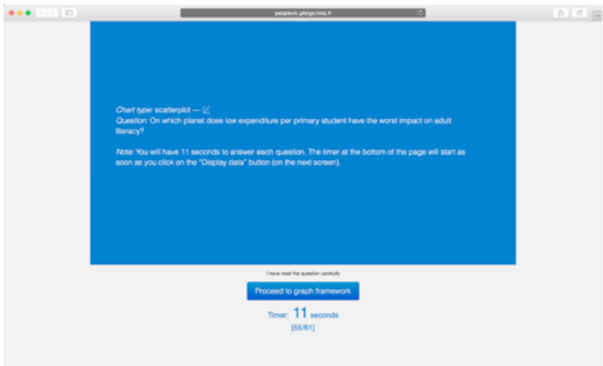


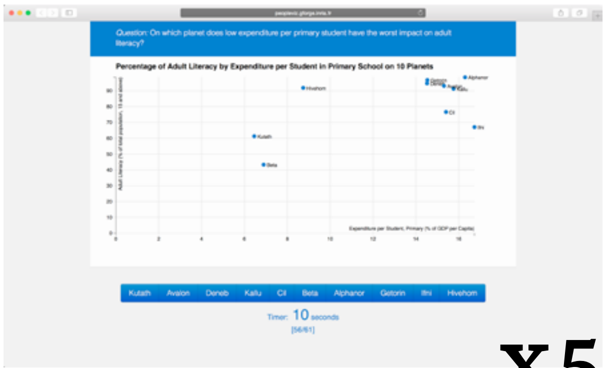
Chart type: scatterplot —
Question: On which planet does low expenditure per primary student have the worst impact on adult literacy?
Note: You will have 11 seconds to answer each question. The timer at the bottom of the page will start as soon as you click on the "Display data" button (on the next screen).
Proceed to graph framework
Timer: 11 seconds
[55/61]


Question: On which planet does low expenditure per primary student have the worst impact on adult literacy?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds
[55/61]

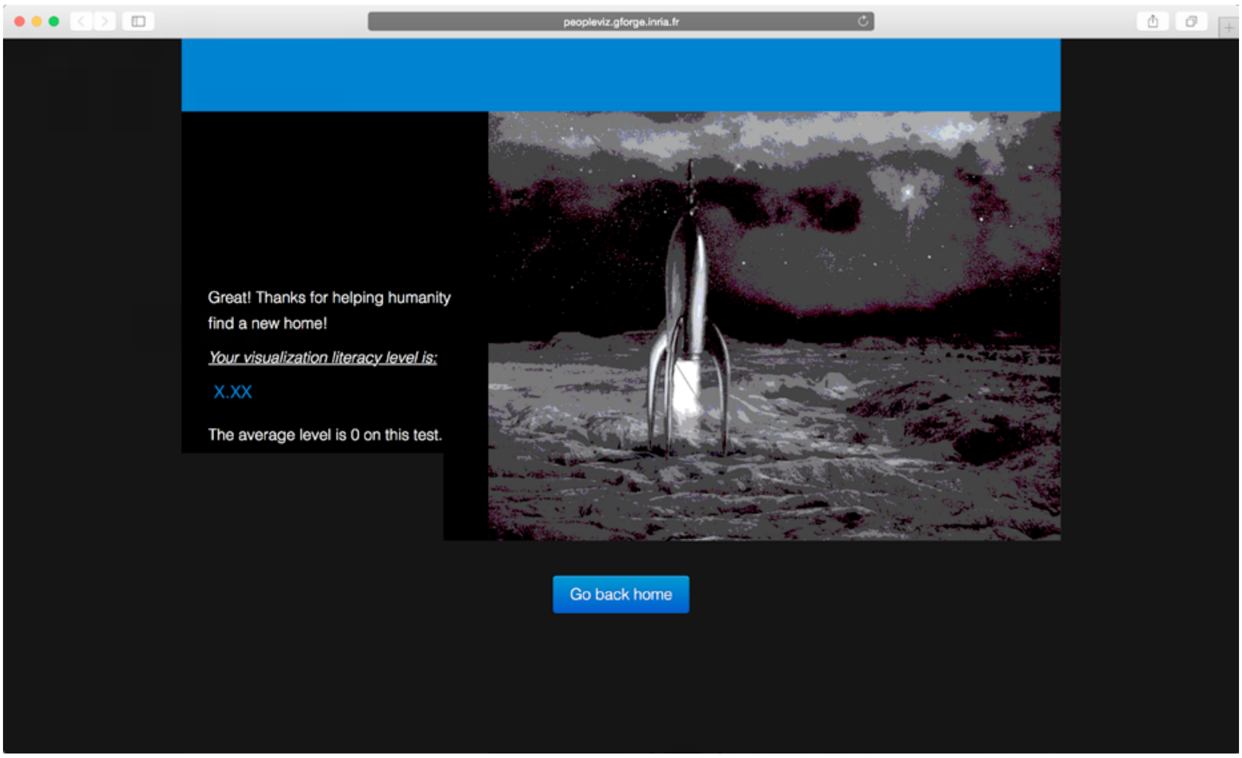

Question: On which planet does low expenditure per primary student have the worst impact on adult literacy?
Table: Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Demetria Appelax Avalon Samstead Maag Aerie Storus Rann Elaris Scylla
Timer: 10 seconds
[55/61]

x1


Question: On which planet does low expenditure per primary student have the worst impact on adult literacy?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Display data
Timer: 11 seconds
[56/61]


Question: On which planet does low expenditure per primary student have the worst impact on adult literacy?
Percentage of Adult Literacy by Expenditure per Student in Primary School on 10 Planets
Kutath Avalon Deneb Kallu Cil Beta Alphanor Getorin Ifni Hvehom
Timer: 10 seconds
[56/61]

x5


peopleviz.gforge.inria.fr
Great! Thanks for helping humanity find a new home!
Your visualization literacy level is:
X.XX
The average level is 0 on this test.
Go back home

Appendix I

Inria: Technologies Appliquées

The visualization can be found at http://peopleviz.gforge.inria.fr/trunk/inria_vis/.


Inria
INVENTEURS DU MONDE NUMÉRIQUE
RAPPORT ANNUEL 2012
LE NUMÉRIQUE,
ENJEU SOCIAL

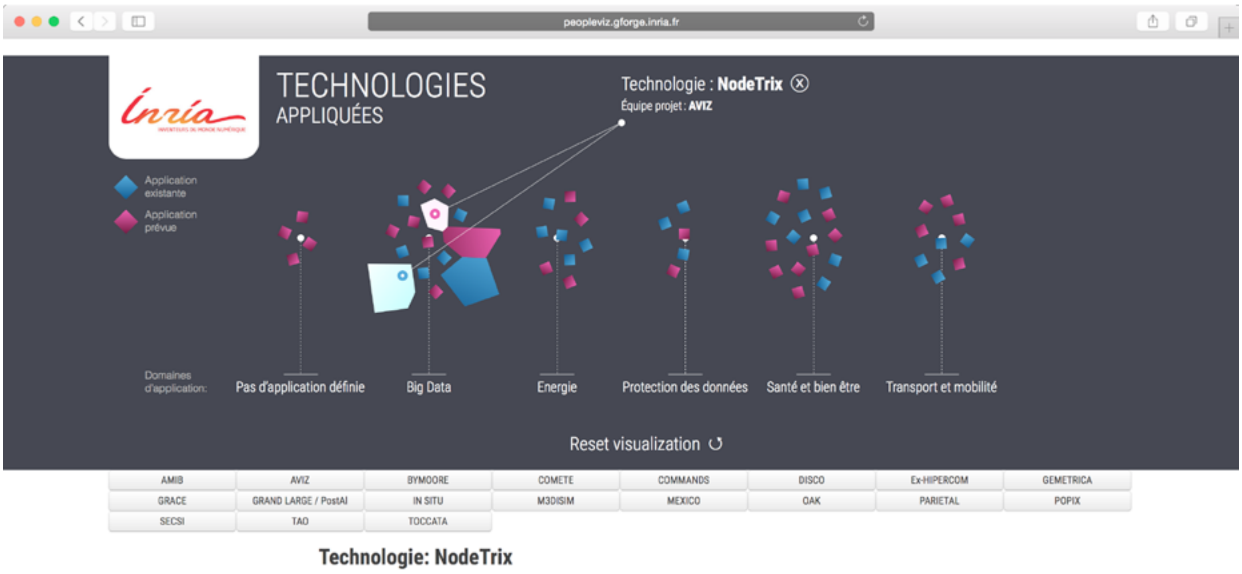


FIGURE I.1: **Inria: Technologies Appliquées.** This visualization I created presents different Inria research projects that have been transferred to industrial applications. It uses free visual variables to create the same kind of 'imagery' as the institutions' 2012 annual report, *i.e,* the floating particles, the way the logo is inserted, a similar font, *etc.* This provides a consistent 'look and feel' between the documents.

Appendix J

Testing Interaction Propensity: Experiment 1

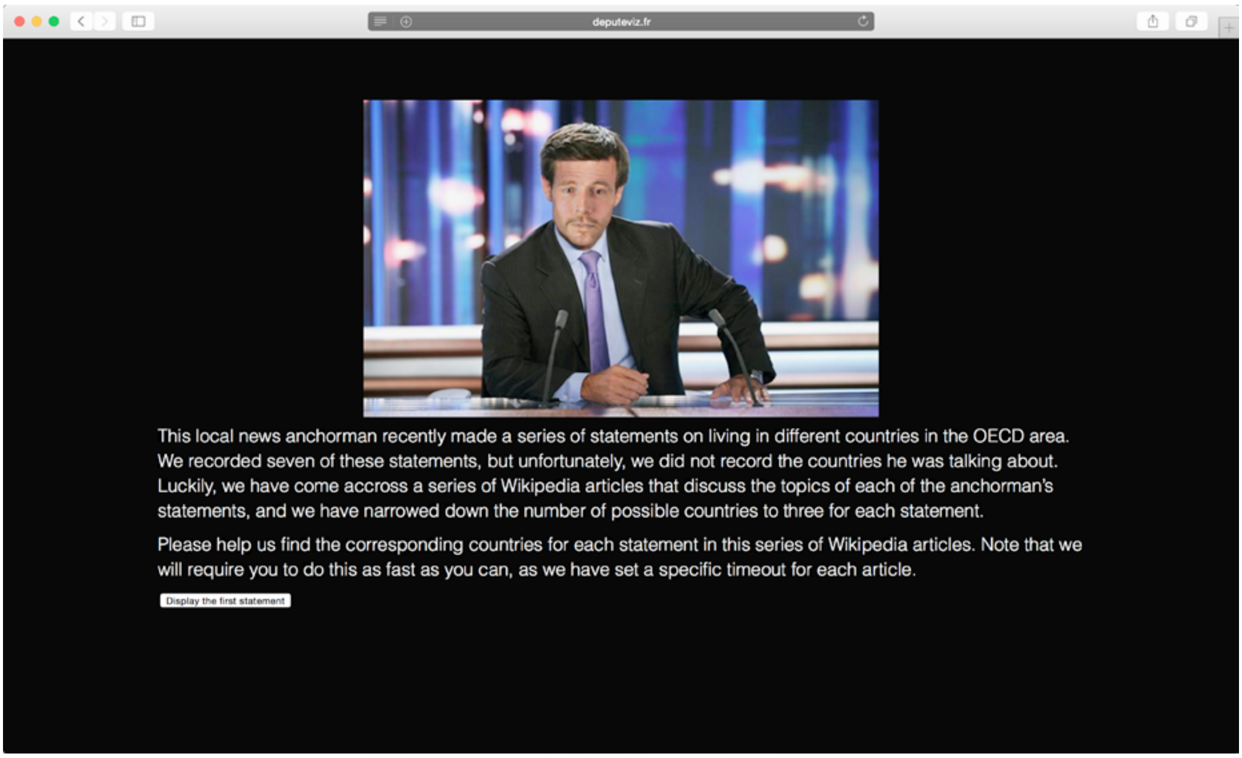
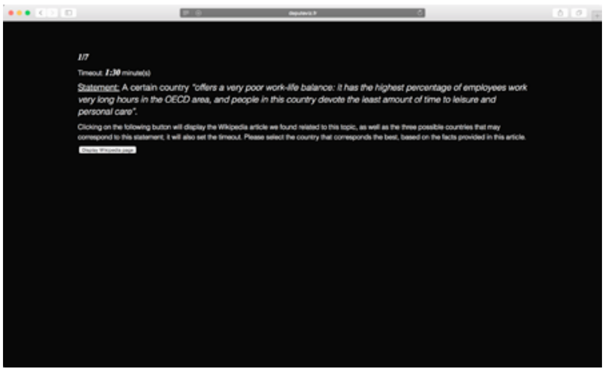
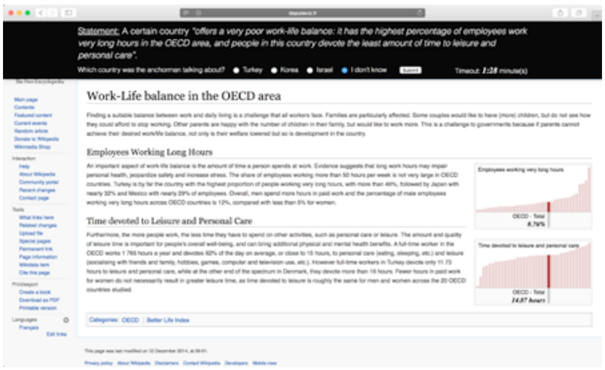
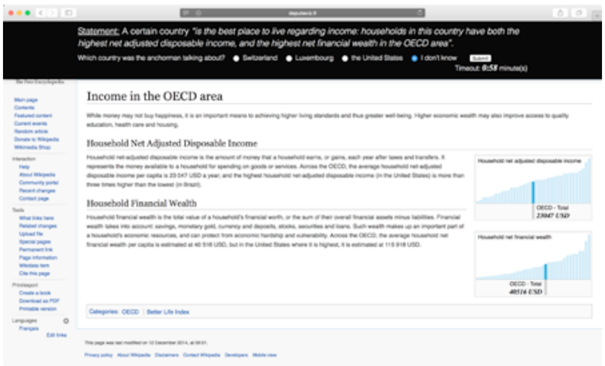


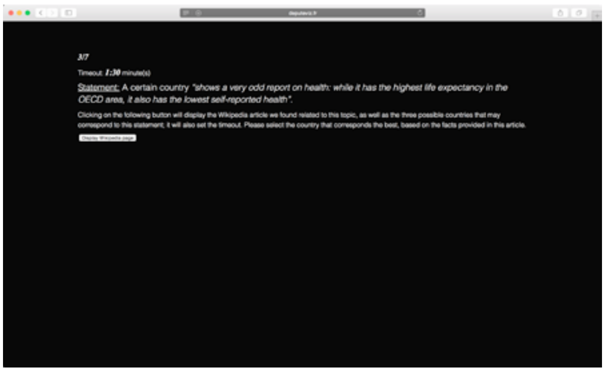
FIGURE J.1: The 7 articles used in Experiment 1 and their sequencing.



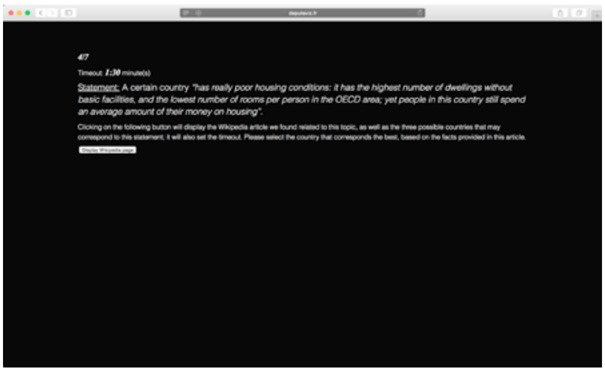


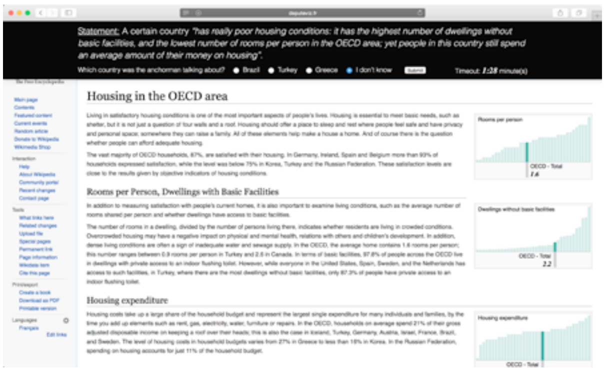

2/7
Timeout: 1:00 minute(s)
Statement: A certain country "is the best place to live regarding income: households in this country have both the highest net adjusted disposable income, and the highest net financial wealth in the OECD area".


Income in the OECD area
Household Net Adjusted Disposable Income
Household Financial Wealth

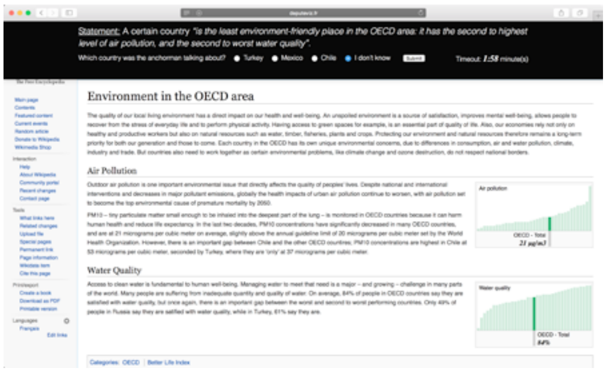

3/7
Timeout: 1:30 minute(s)
Statement: A certain country "shows a very odd report on health: while it has the highest life expectancy in the OECD area, it also has the lowest self-reported health".

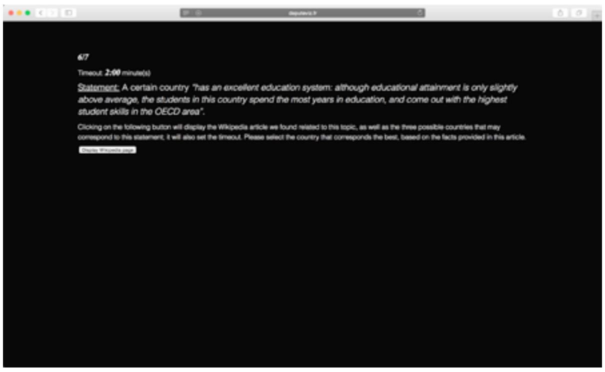

Health in the OECD area
Life Expectancy
Self-Reported Health

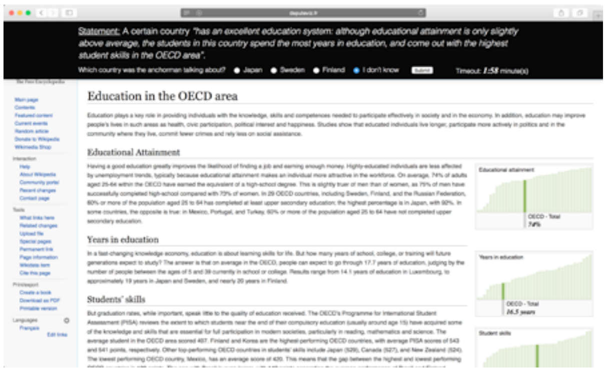

4/7
Timeout: 1:30 minute(s)
Statement: A certain country "has really poor housing conditions: it has the highest number of dwellings without basic facilities, and the lowest number of rooms per person in the OECD area; yet people in this country still spend an average amount of their money on housing".


Housing in the OECD area
Rooms per Person, Dwellings with Basic Facilities
Housing expenditure

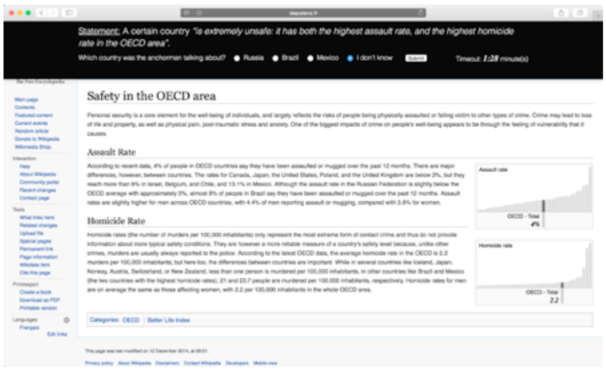

5/7
Timeout: 2:00 minute(s)
Statement: A certain country "is the least environment-friendly place in the OECD area: it has the second to highest level of air pollution, and the second to worst water quality".


Statement: A certain country "is the least environment-friendly place in the OECD area: it has the second to highest level of air pollution, and the second to worst water quality".
Environment in the OECD area
Air Pollution
Water Quality


6/7
Timeout: 2:00 minute(s)
Statement: A certain country "has an excellent education system: although educational attainment is only slightly above average, the students in this country spend the most years in education, and come out with the highest student skills in the OECD area".


Statement: A certain country "has an excellent education system: although educational attainment is only slightly above average, the students in this country spend the most years in education, and come out with the highest student skills in the OECD area".
Education in the OECD area
Educational Attainment
Years in education
Students' skills


7/7
Timeout: 1:30 minute(s)
Statement: A certain country "is extremely unsafe: it has both the highest assault rate, and the highest homicide rate in the OECD area".


Statement: A certain country "is extremely unsafe: it has both the highest assault rate, and the highest homicide rate in the OECD area".
Safety in the OECD area
Assault Rate
Homicide Rate

Appendix K

Testing Interaction Propensity: Experiment 2

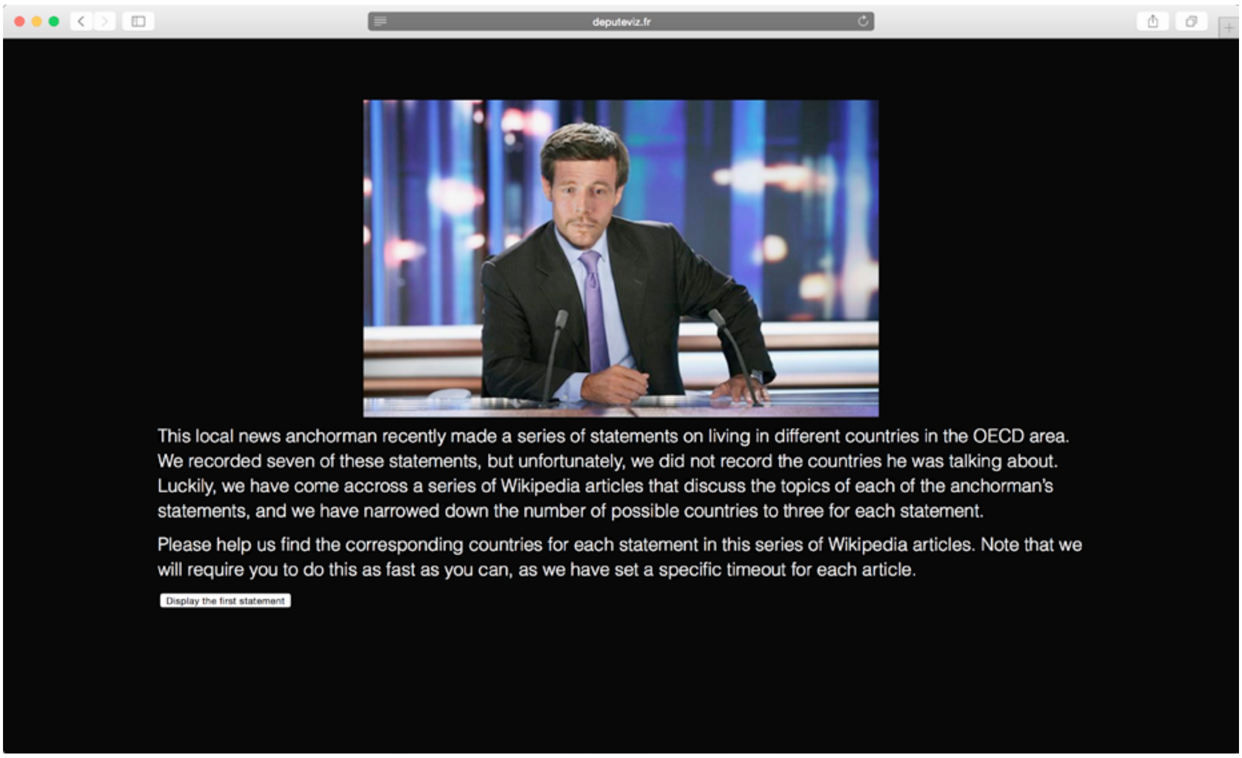
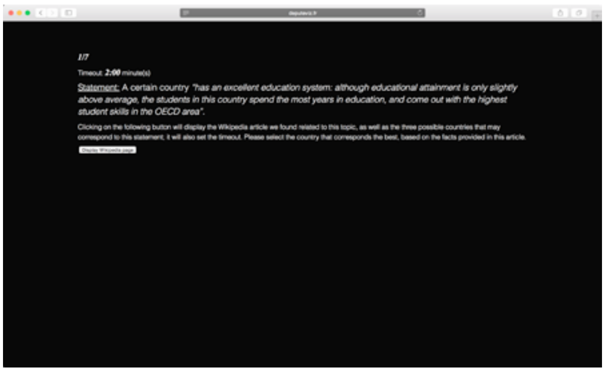
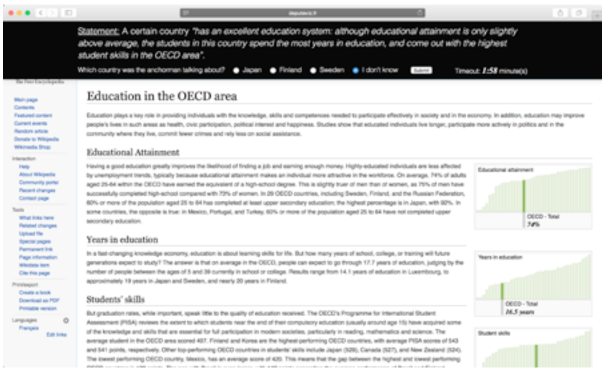
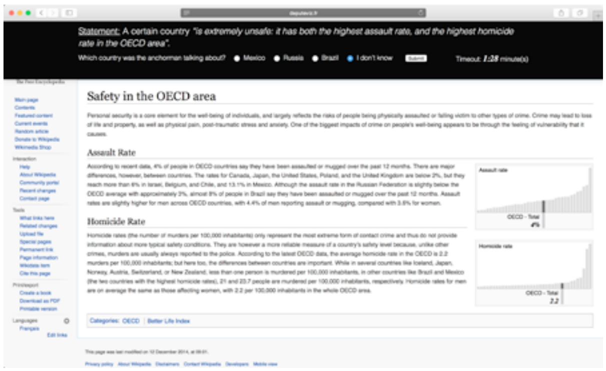


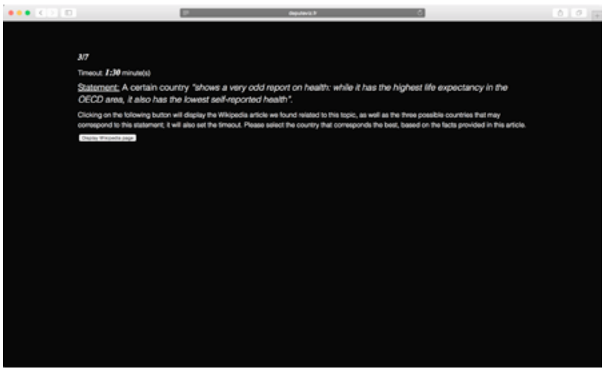
FIGURE K.1: The 7 articles used in Experiment 2 and their sequencing.



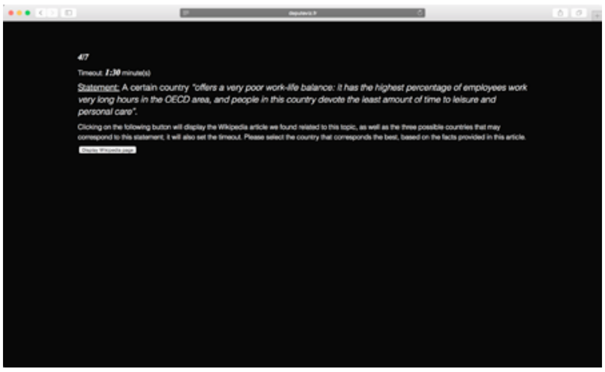



2/7
Timeout: 1:30 minute(s)
Statement: A certain country "is extremely unsafe: it has both the highest assault rate, and the highest homicide rate in the OECD area".
Clicking on the following button will display the Wikipedia article we found related to this topic, as well as the three possible countries that may correspond to this statement; it will also set the timeout. Please select the country that corresponds the best, based on the facts provided in this article.


Statement: A certain country "is extremely unsafe: it has both the highest assault rate, and the highest homicide rate in the OECD area".
Which country was the anchorman talking about? Mexico Russia Brazil I don't know
Timeout: 1:28 minute(s)
Safety in the OECD area
Assault Rate
Homicide Rate

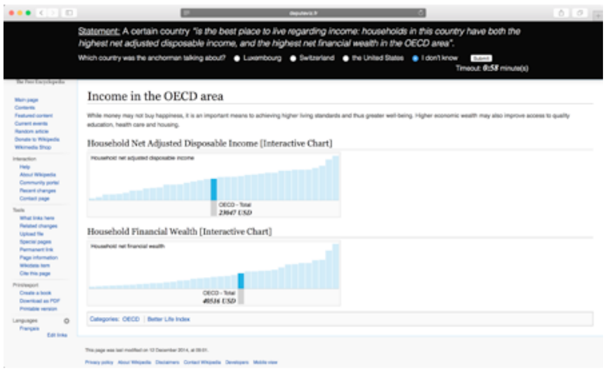

3/7
Timeout: 1:30 minute(s)
Statement: A certain country "shows a very odd report on health: while it has the highest life expectancy in the OECD area, it also has the lowest self-reported health".
Clicking on the following button will display the Wikipedia article we found related to this topic, as well as the three possible countries that may correspond to this statement; it will also set the timeout. Please select the country that corresponds the best, based on the facts provided in this article.


Statement: A certain country "shows a very odd report on health: while it has the highest life expectancy in the OECD area, it also has the lowest self-reported health".
Which country was the anchorman talking about? Switzerland Iceland Japan I don't know
Timeout: 1:19 minute(s)
Health in the OECD area
Life Expectancy [Interactive Chart]
Self-Reported Health [Interactive Chart]

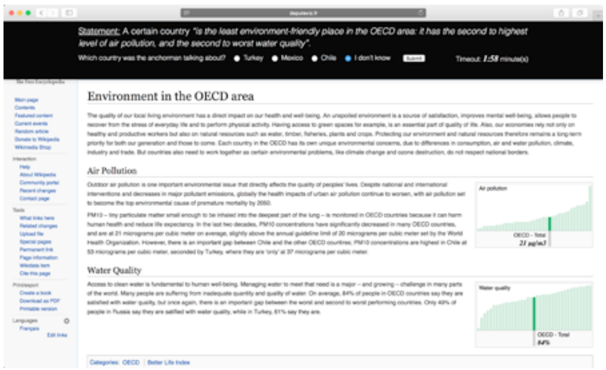

4/7
Timeout: 1:30 minute(s)
Statement: A certain country "offers a very poor work-life balance: it has the highest percentage of employees work very long hours in the OECD area, and people in this country devote the least amount of time to leisure and personal care".
Clicking on the following button will display the Wikipedia article we found related to this topic, as well as the three possible countries that may correspond to this statement; it will also set the timeout. Please select the country that corresponds the best, based on the facts provided in this article.

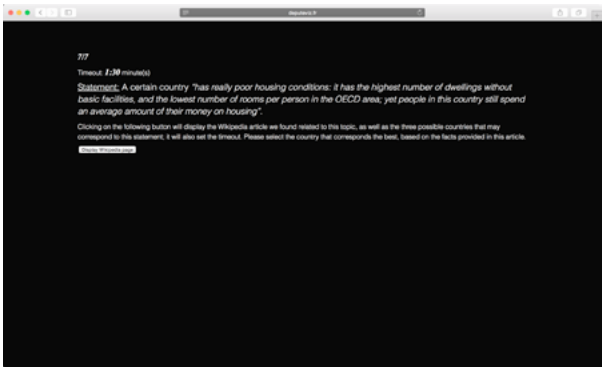

Statement: A certain country "offers a very poor work-life balance: it has the highest percentage of employees work very long hours in the OECD area, and people in this country devote the least amount of time to leisure and personal care".
Which country was the anchorman talking about? Turkey Korea Israel I don't know
Timeout: 1:28 minute(s)
Work-Life balance in the OECD area
Employees Working Long Hours [Interactive Chart]
Time devoted to Leisure and Personal Care [Interactive Chart]

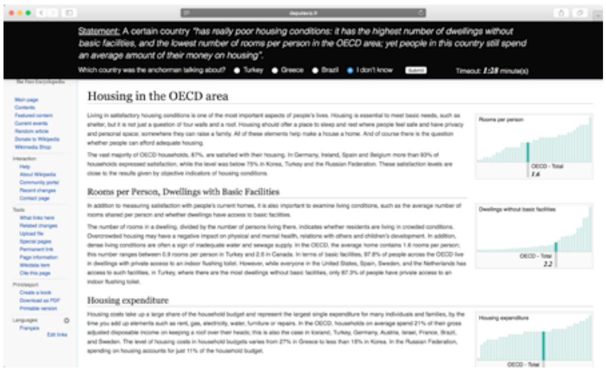

5/7
Timeout: 1:00 minute(s)
Statement: A certain country "is the best place to live regarding income: households in this country have both the highest net adjusted disposable income, and the highest net financial wealth in the OECD area".


Statement: A certain country "is the best place to live regarding income: households in this country have both the highest net adjusted disposable income, and the highest net financial wealth in the OECD area".
Which country was the anchorman talking about? Luxembourg Switzerland the United States I don't know
Income in the OECD area
Household Net Adjusted Disposable Income [Interactive Chart]
Household Financial Wealth [Interactive Chart]


6/7
Timeout: 2:00 minute(s)
Statement: A certain country "is the least environment-friendly place in the OECD area: it has the second to highest level of air pollution, and the second to worst water quality".


Statement: A certain country "is the least environment-friendly place in the OECD area: it has the second to highest level of air pollution, and the second to worst water quality".
Which country was the anchorman talking about? Turkey Mexico Chile I don't know
Environment in the OECD area
Air Pollution
Water Quality


7/7
Timeout: 1:30 minute(s)
Statement: A certain country "has really poor housing conditions: it has the highest number of dwellings without basic facilities, and the lowest number of rooms per person in the OECD area; yet people in this country still spend an average amount of their money on housing".


Statement: A certain country "has really poor housing conditions: it has the highest number of dwellings without basic facilities, and the lowest number of rooms per person in the OECD area; yet people in this country still spend an average amount of their money on housing".
Which country was the anchorman talking about? Turkey Greece Brazil I don't know
Housing in the OECD area
Rooms per Person, Dwellings with Basic Facilities
Housing expenditure

Appendix L

Testing Interaction Propensity: Experiment 3

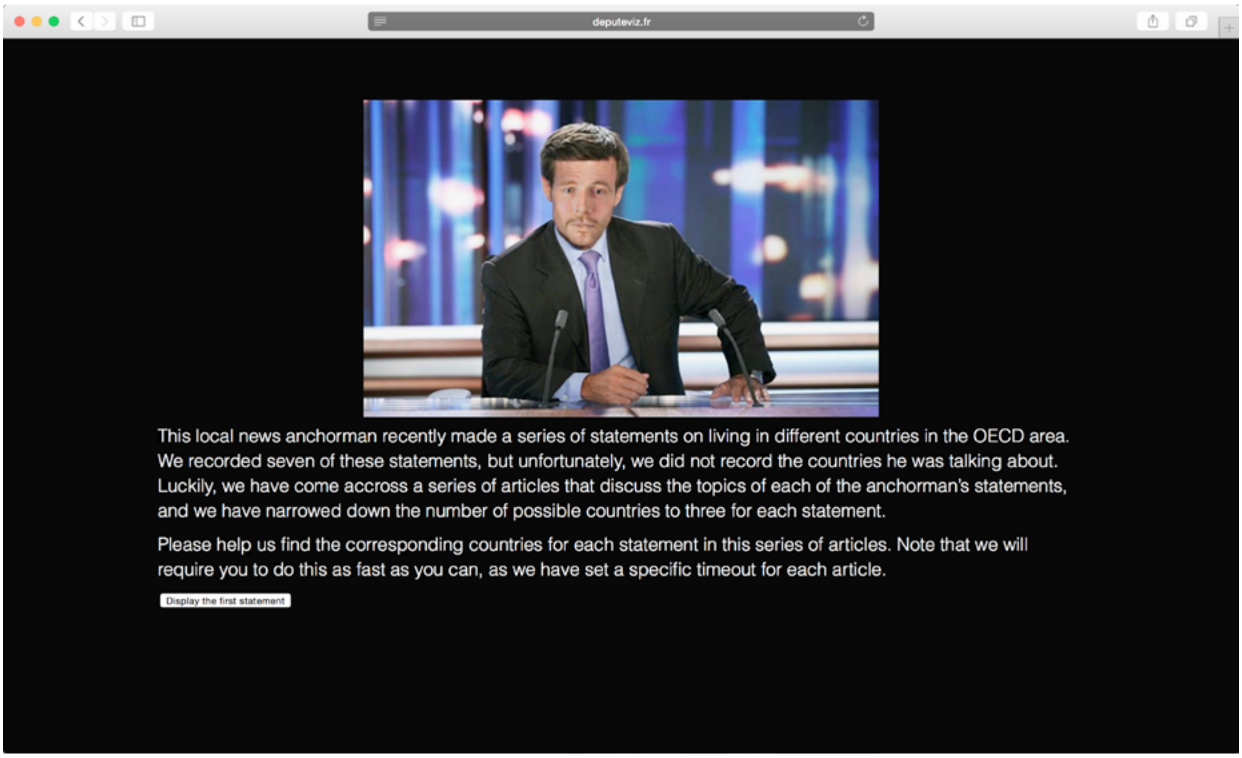
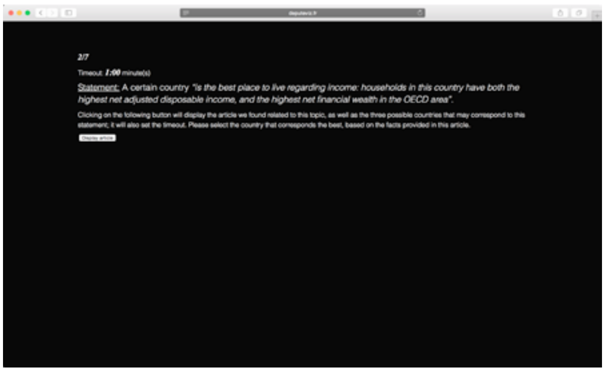
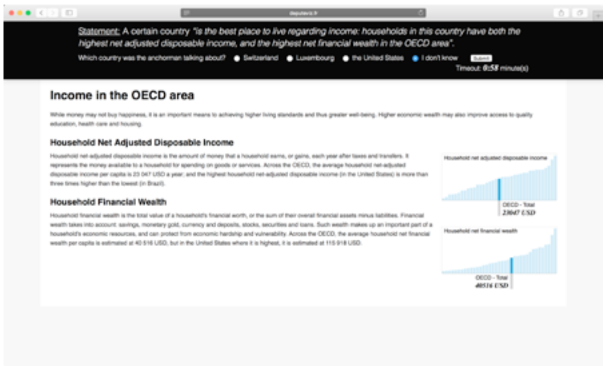
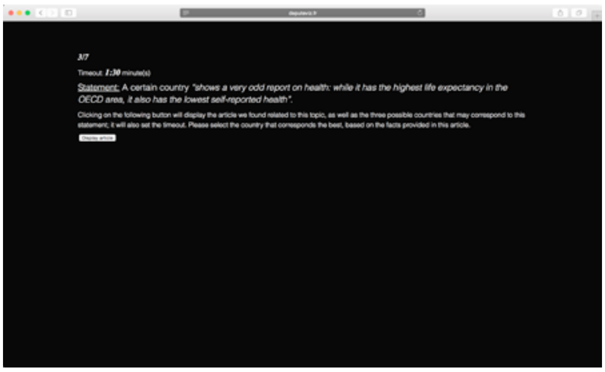
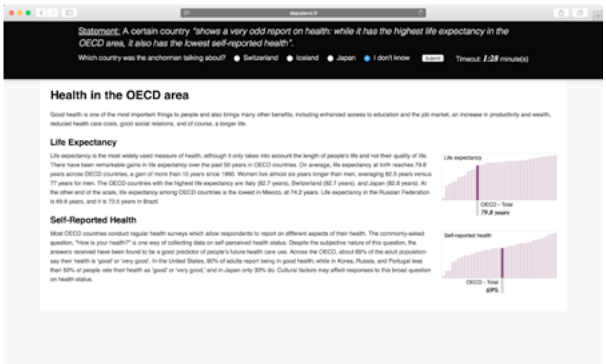


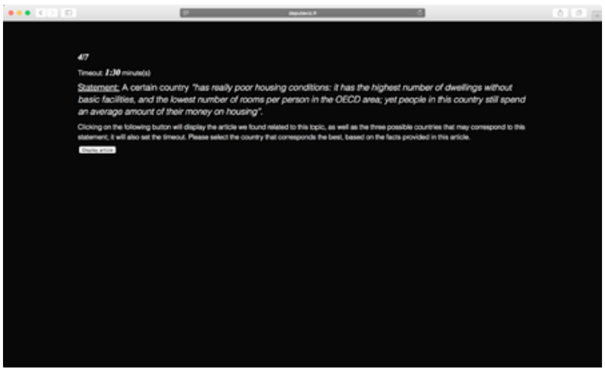
FIGURE L.1: The 7 articles used in Experiment 3 and their sequencing.

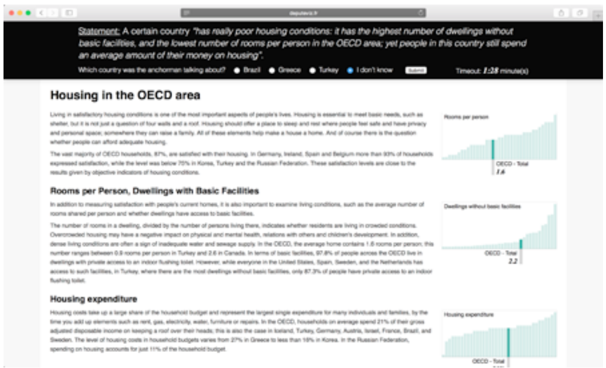

2/7
Statement: A certain country "is the best place to live regarding income: households in this country have both the highest net adjusted disposable income, and the highest net financial wealth in the OECD area".

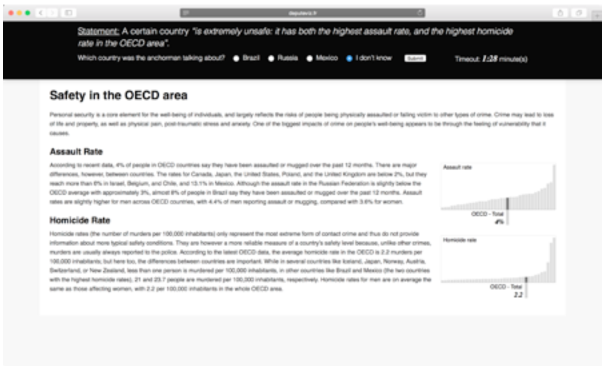

Statement: A certain country "is the best place to live regarding income: households in this country have both the highest net adjusted disposable income, and the highest net financial wealth in the OECD area".
Income in the OECD area
Household Net Adjusted Disposable Income
Household Financial Wealth

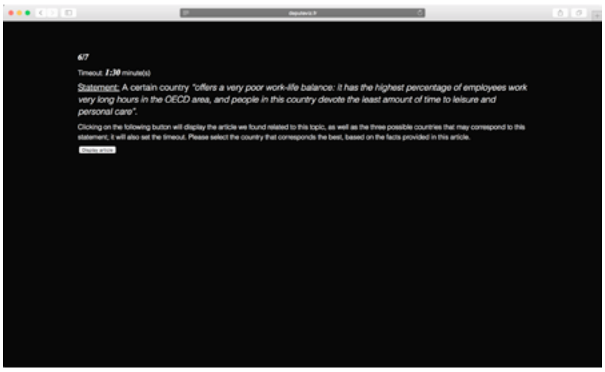

3/7
Statement: A certain country "shows a very odd report on health: while it has the highest life expectancy in the OECD area, it also has the lowest self-reported health".

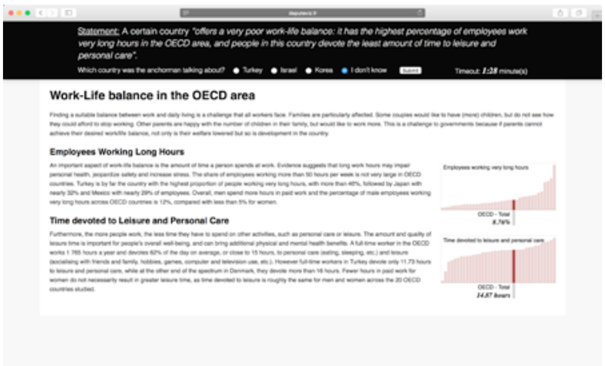

Statement: A certain country "shows a very odd report on health: while it has the highest life expectancy in the OECD area, it also has the lowest self-reported health".
Health in the OECD area
Life Expectancy
Self-Reported Health

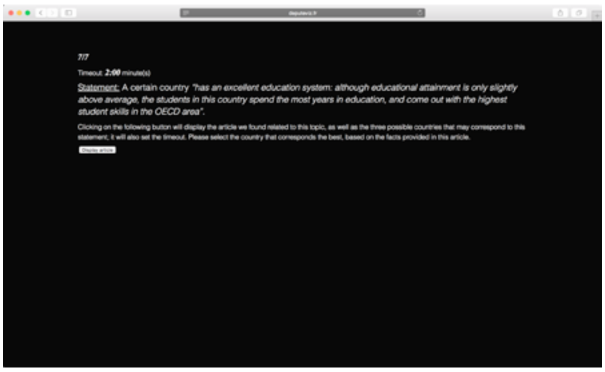

4/7
Statement: A certain country "has really poor housing conditions: it has the highest number of dwellings without basic facilities, and the lowest number of rooms per person in the OECD area; yet people in this country still spend an average amount of their money on housing".


Statement: A certain country "has really poor housing conditions: it has the highest number of dwellings without basic facilities, and the lowest number of rooms per person in the OECD area; yet people in this country still spend an average amount of their money on housing".
Housing in the OECD area
Rooms per Person, Dwellings with Basic Facilities
Housing expenditure









Appendix M

The CO2 Pollution Explorer

The visualization can be found at http://peopleviz.gforge.inria.fr/trunk/data_blog/environment/co2/index_storytelling.php.

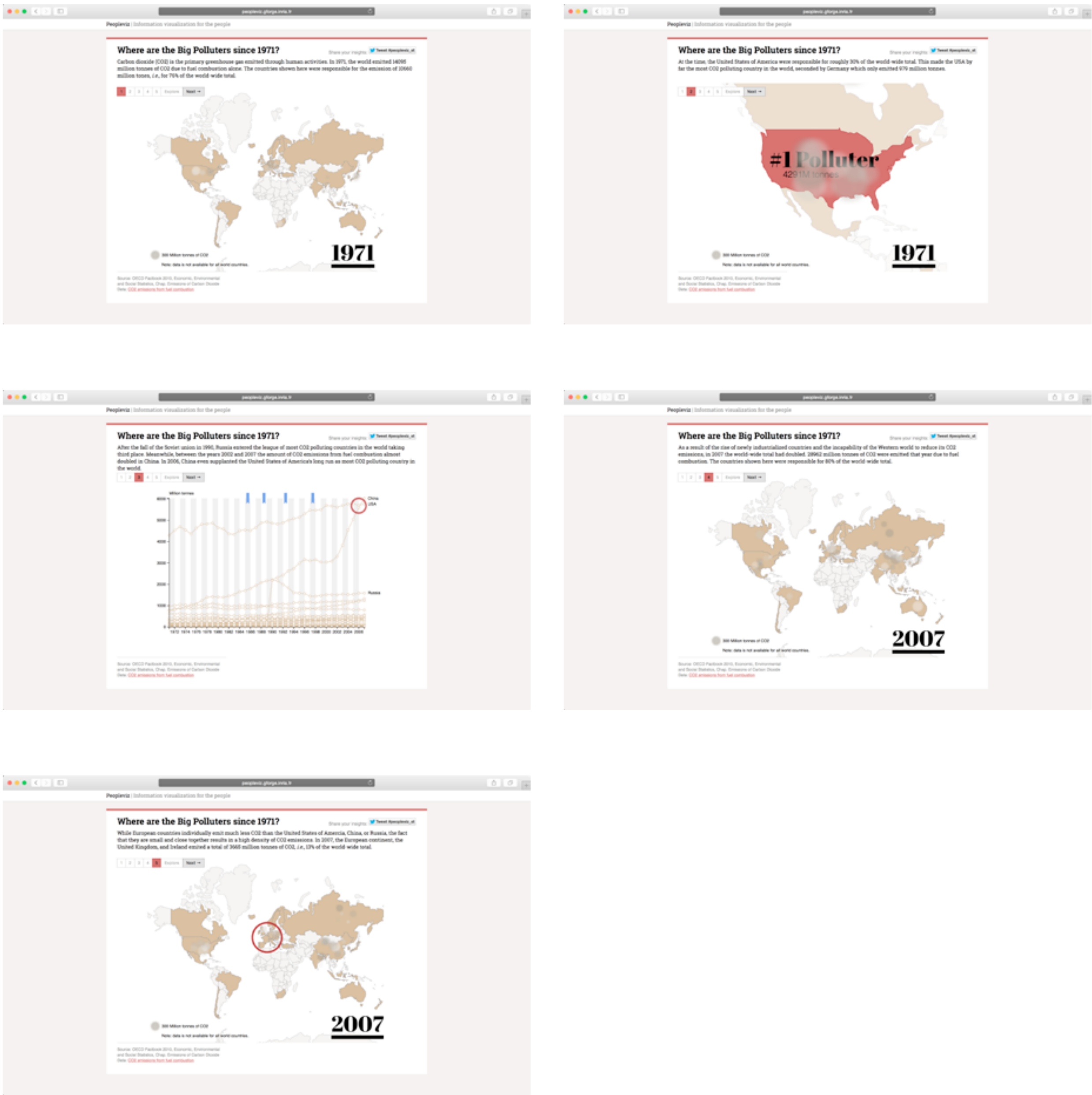
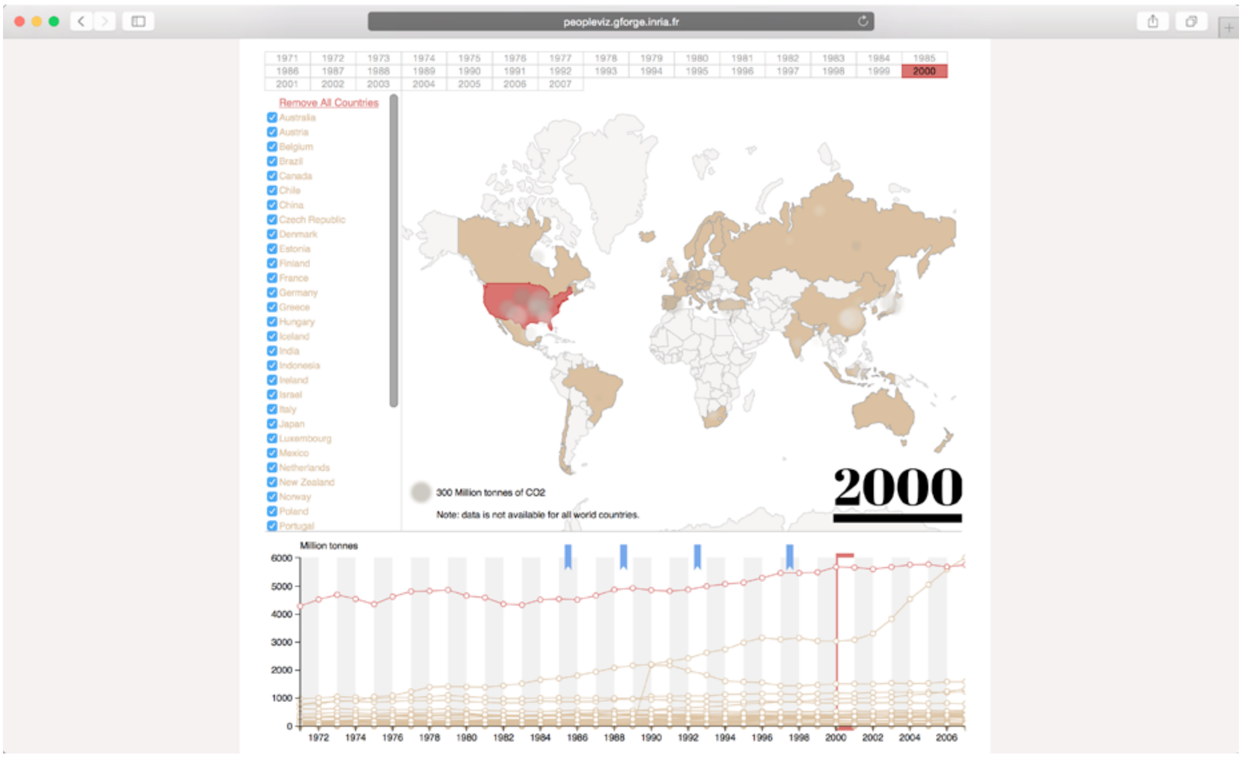


FIGURE M.1: The 6 sections of the ST version of the *CO2 Pollution Explorer* (including the Explore section—next page).



peopleviz.gforge.inria.fr
Remove All Countries
Australia
Austria
Belgium
Brazil
Canada
Chile
China
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hungary
Iceland
India
Indonesia
Ireland
Israel
Italy
Japan
Luxembourg
Mexico
Netherlands
New Zealand
Norway
Poland
Portugal
300 Million tonnes of CO2
Note: data is not available for all world countries.
2000
Million tonnes

Appendix N

The Economic Return on Education Explorer

The visualization can be found at http://peopleviz.gforge.inria.fr/trunk/mediaviz_EN/.

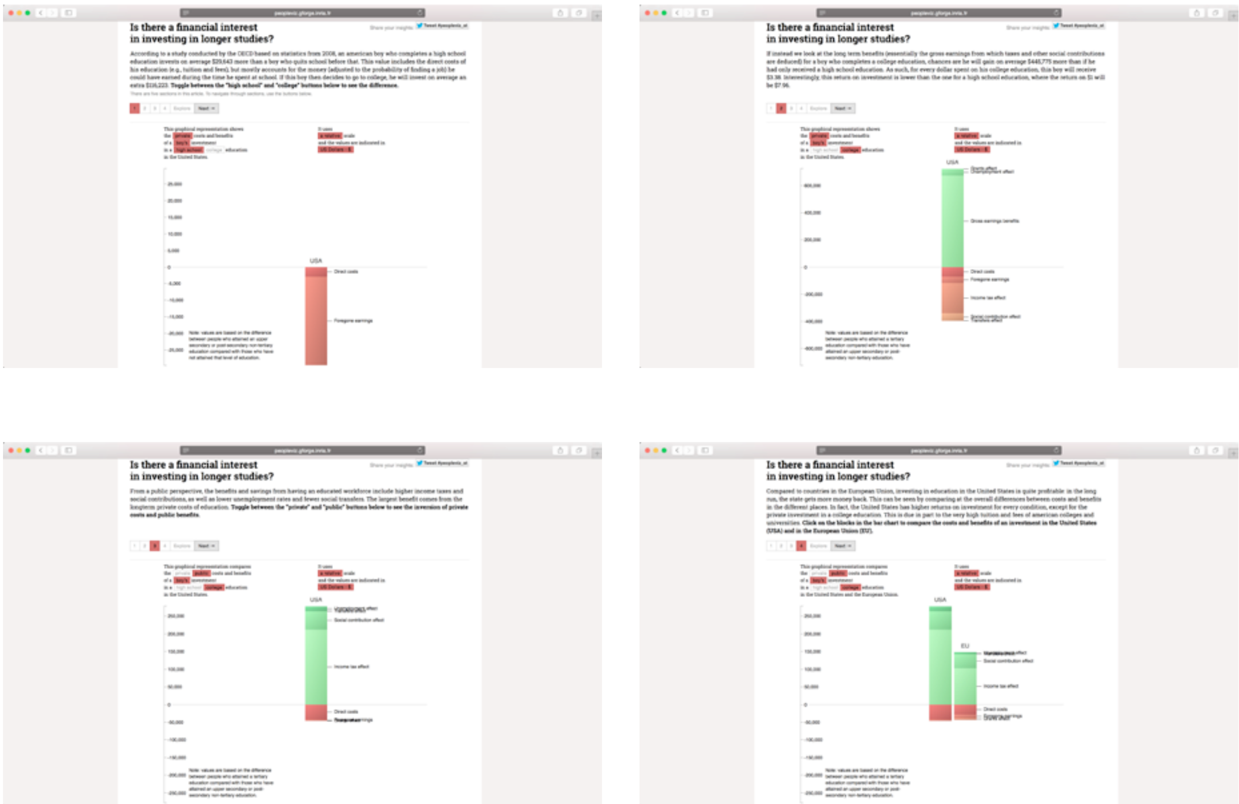
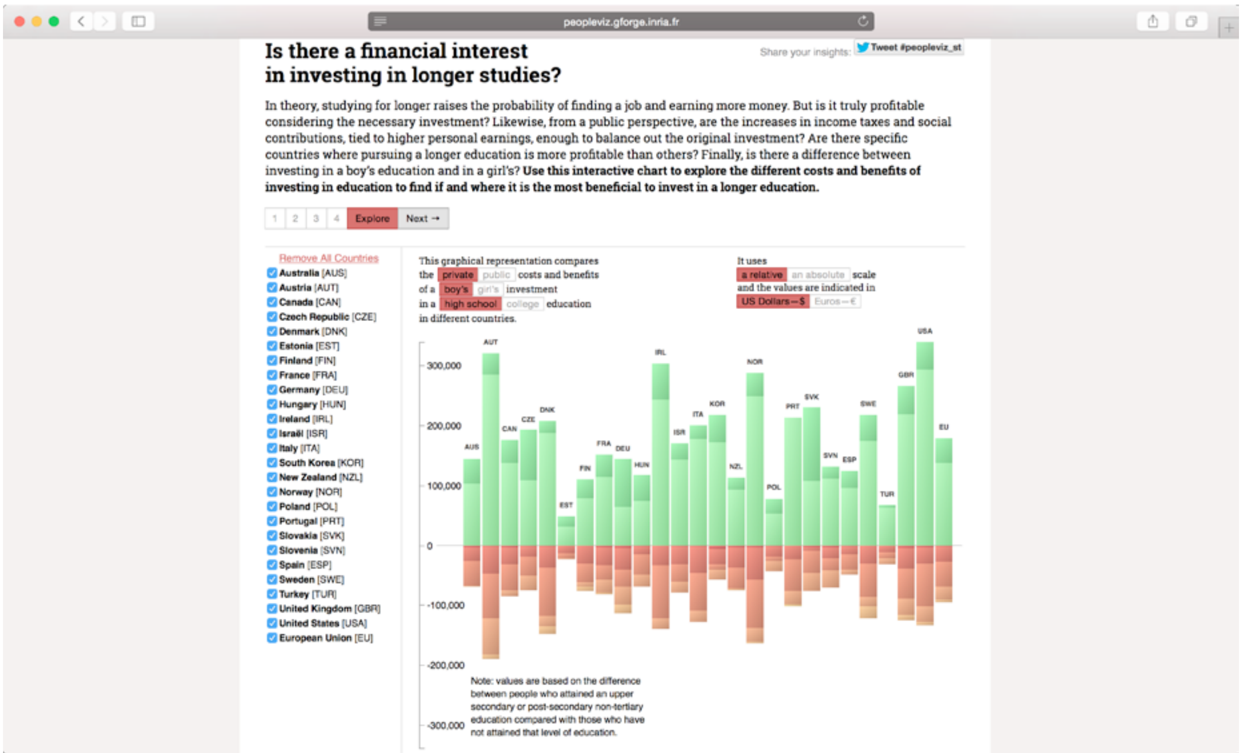


FIGURE N.1: The 5 sections of the ST version of the *Economic Return on Eductation Explorer* (including the Explore section—next page).


peopleviz.gforge.inria.fr
Is there a financial interest in investing in longer studies?
Share your insights: Tweet #peopleviz_st
In theory, studying for longer raises the probability of finding a job and earning more money. But is it truly profitable considering the necessary investment? Likewise, from a public perspective, are the increases in income taxes and social contributions, tied to higher personal earnings, enough to balance out the original investment? Are there specific countries where pursuing a longer education is more profitable than others? Finally, is there a difference between investing in a boy's education and in a girl's? Use this interactive chart to explore the different costs and benefits of investing in education to find if and where it is the most beneficial to invest in a longer education.
1 2 3 4 Explore Next →
Remove All Countries
Australia [AUS]
Austria [AUT]
Canada [CAN]
Czech Republic [CZE]
Denmark [DNK]
Estonia [EST]
Finland [FIN]
France [FRA]
Germany [DEU]
Hungary [HUN]
Ireland [IRL]
Israël [ISR]
Italy [ITA]
South Korea [KOR]
New Zealand [NZL]
Norway [NOR]
Poland [POL]
Portugal [PRT]
Slovakia [SVK]
Slovenia [SVN]
Spain [ESP]
Sweden [SWE]
Turkey [TUR]
United Kingdom [GBR]
United States [USA]
European Union [EU]
This graphical representation compares the private public costs and benefits of a boy's girl's investment in a high school college education in different countries.
It uses a relative an absolute scale and the values are indicated in US Dollars—$ Euros—€
300,000
200,000
100,000
0
-100,000
-200,000
-300,000
AUS AUT CAN CZE DNK EST FIN FRA DEU HUN IRL ISR ITA KOR NZL NOR POL PRT SVK SVN ESP SWE TUR GBR USA EU
Note: values are based on the difference between people who attained an upper secondary or post-secondary non-tertiary education compared with those who have not attained that level of education.

Appendix O

The Nuclear Power Grid

The visualization can be found at http://peopleviz.gforge.inria.fr/trunk/data_blog/environment/nuclear/.

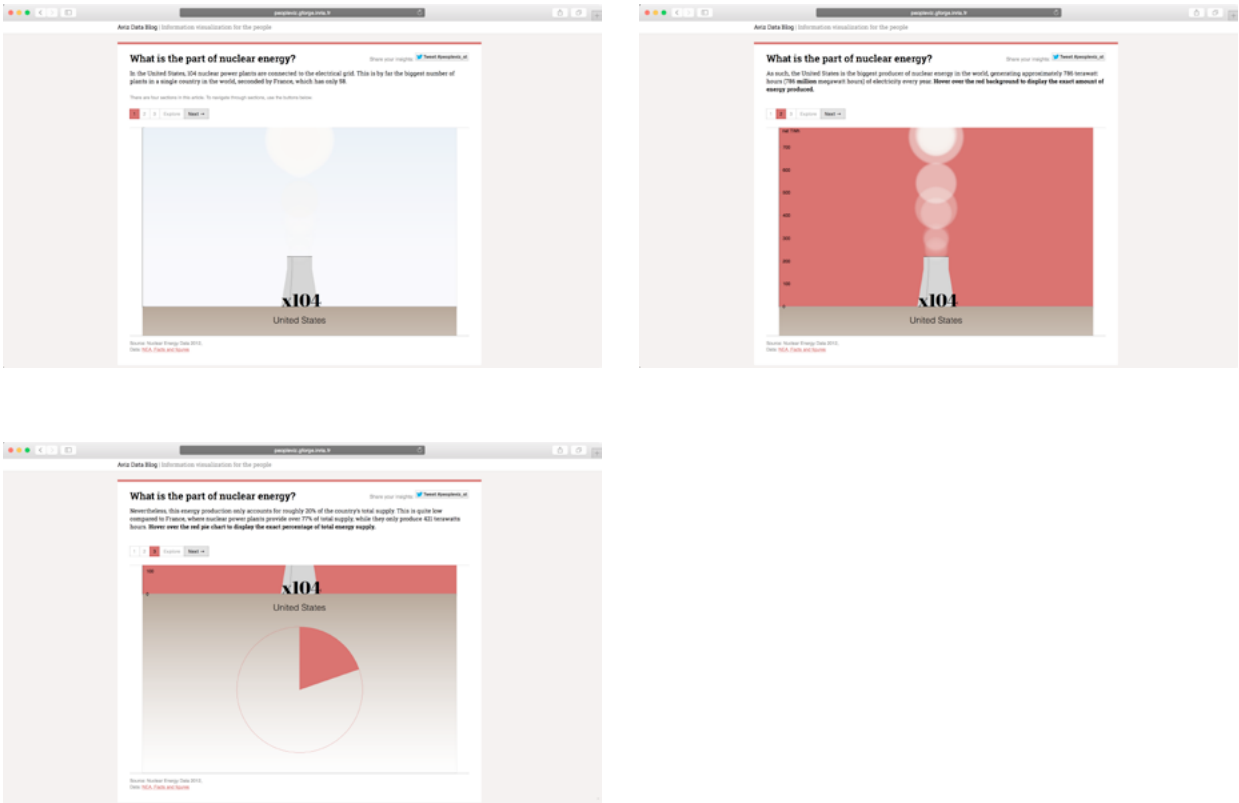
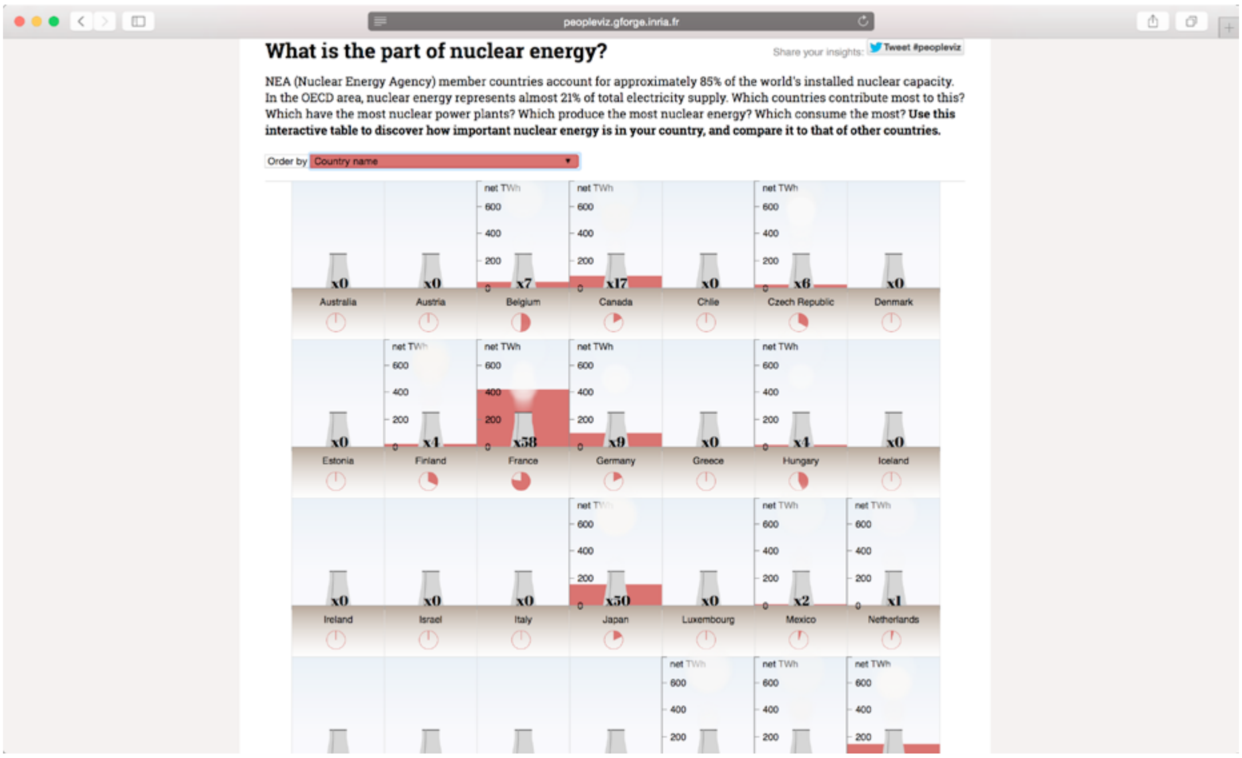


FIGURE O.1: The 4 sections of the ST version of the *Nuclear Power Grid* (including the Explore section—next page).


peopleviz.gforge.inria.fr
What is the part of nuclear energy?
Share your insights: Tweet #peopleviz
NEA (Nuclear Energy Agency) member countries account for approximately 85% of the world's installed nuclear capacity. In the OECD area, nuclear energy represents almost 21% of total electricity supply. Which countries contribute most to this? Which have the most nuclear power plants? Which produce the most nuclear energy? Which consume the most? Use this interactive table to discover how important nuclear energy is in your country, and compare it to that of other countries.
Order by Country name
Australia
Austria
Belgium
Canada
Chile
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hungary
Iceland
Ireland
Israel
Italy
Japan
Luxembourg
Mexico
Netherlands

Appendix P

The French Deputies Explorer

The visualization can be found at http://www.deputeviz.fr/.

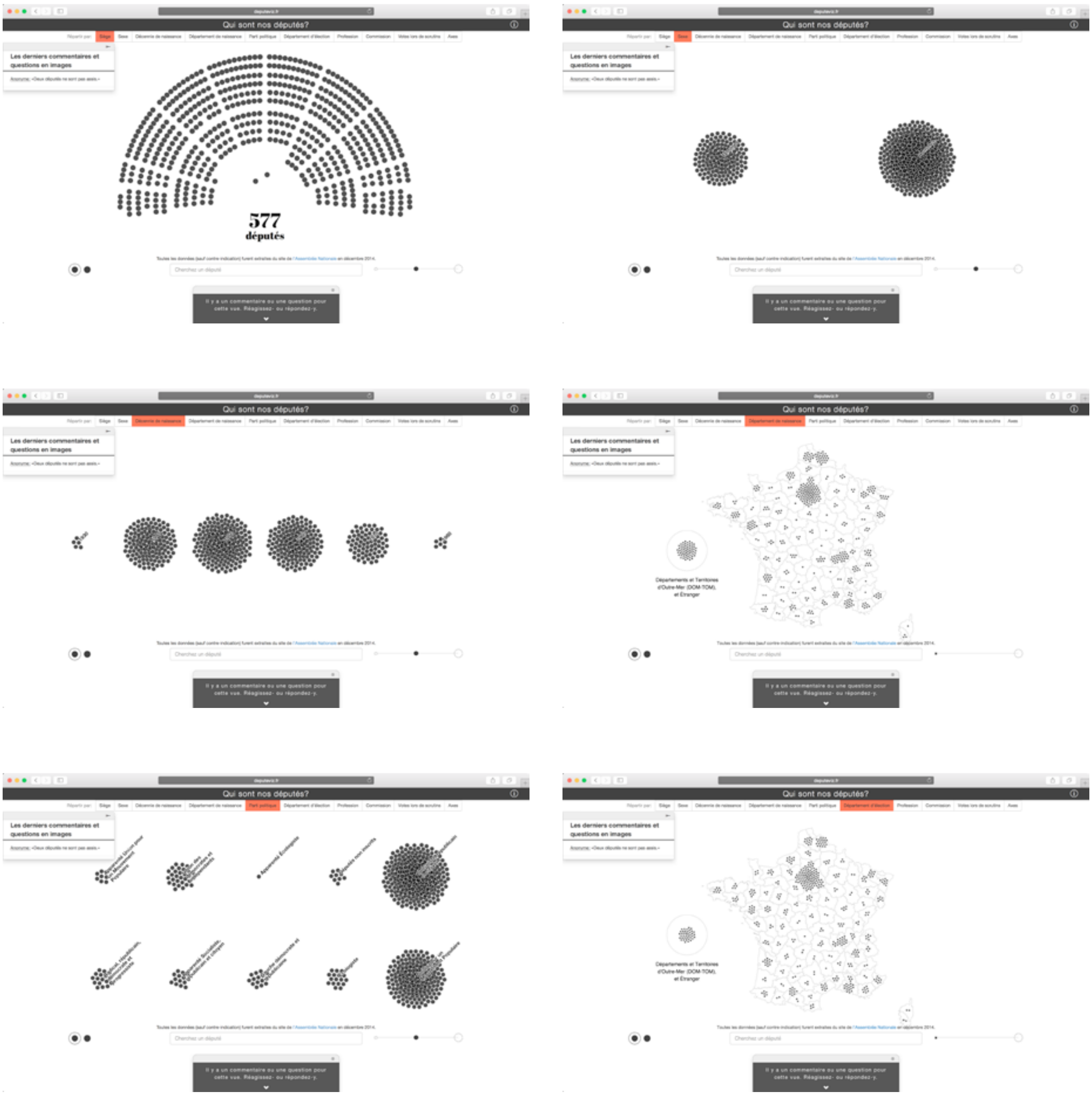
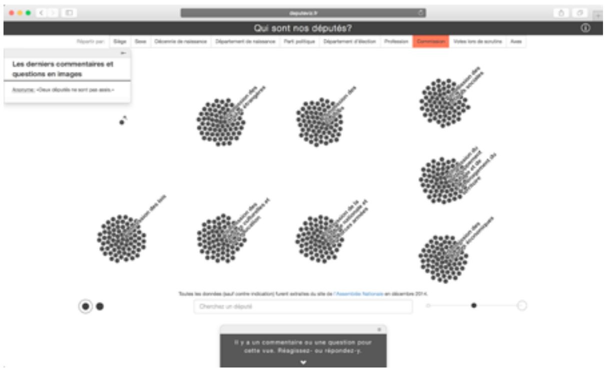
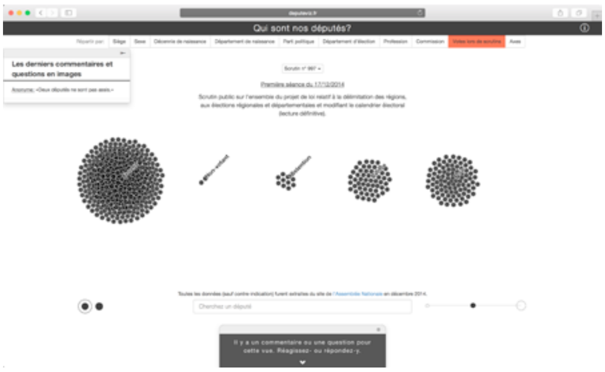
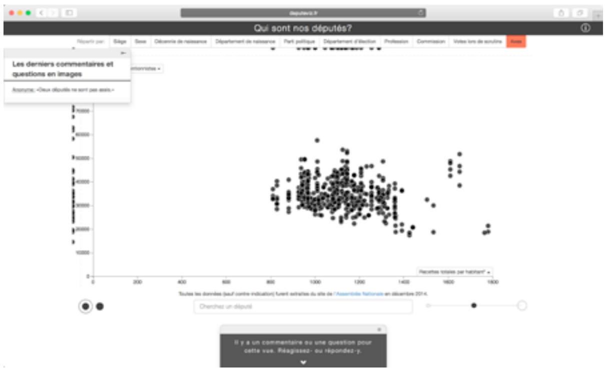


FIGURE P.1: The 10 different views of the *French Deputies Explorer* (Deputeviz—Figure is continued on next page).


Qui sont nos députés?


Qui sont nos députés?


Qui sont nos députés?


Qui sont nos députés?

Bibliography

[1] http://blog.visual.ly/interaction-design-for-data-visualizations/.

[2] http://data.worldbank.org/.

[3] http://datavizblog.com/category/london-1854-cholera-epidemic/.

[4] http://deputeviz.fr/other_apps/SI/wikipedia_template/example.php?t=1&s=1.

[5] http://html5gallery.com/.

[6] http://numericalreasoningtest.org/.

[7] http://opensource.org/docs/osd.

[8] http://peopleviz.gforge.inria.fr/trunk/data_blog/environment/co2/index_storytelling.php.

[9] http://peopleviz.gforge.inria.fr/trunk/SI_cues/.

[10] http://razor.occams.info/pubdocs/opendataciviccapital. html.

[11] http://readwrite.com/2011/10/02/cartoon_do_i_have_to_draw_you_a_picture.

[12] http://stats.oecd.org/Index.aspx?lang=en.

[13] http://stuartroseman.com/post/619953720/out-with-the-old-business-in-with-the-new.

[14] http://tiny.cc/ujgjdx.

[15] http://tinyurl.com/CommuniqueGouvernement.

[16] http://tinyurl.com/OpenGovAct.

[17] http://tinyurl.com/SourceData01.

[18] http://www.aviz.fr/Research/ValueOfInfoVis.

[19] http://www.aviz.fr/~bbach/vis25timeline/.

[20] http://www.datavis.ca/milestones/index.php?group=1900%2B&mid=ms176.

[21] http://www.economist.com/node/10278643.

[22] http://www.education.gov.uk/schools/careers/traininganddevelopment/professional/b00211213/numeracy/areas/barcharts.

[23] http://www.google.com/design/spec/material-design/introduction.html.

[24] http://www.google.com/publicdata/directory.

[25] http://www.jobtestprep.co.uk/jobtestprep/testPlayer.aspx?localeSetting=en-GB&sLanguage=en-GB&&test_id=free_num_advanced&isSequence= 0IGNORE&testIndex=0&mode=timed&skinID=9& template=FreeTestTemplate&endURL=signout.aspx%3fReturnURL%3d..%2fnumerical-reasoning-IGNOREtest&endURLName=Back+to+JobTestPrep&cancelURLName= Exit+Test&cancelURL=signout.aspx%3fReturnURL%3d..%2fnumerical-reasoning-test&testTitle=Numerical+ Reasoning+-+Table%2fGraph&layoutype=1&adPageID=&itemCode=FreeNum_GMAT.

[26] http://www.justice.gov/oip/blog/foia-update-freedom-information-act-5-usc-sect-552-amended-public-law-no-104-231-110-stat.

[27] http://www.nngroup.com/articles/how-little-do-users-read/.

[28] http://www.nngroup.com/articles/how-long-do-users-stay-on-web.

[29] http://www.nngroup.com/articles/information-scent/.

[30] http://www.oxforddictionaries.com/definition/english/data.

[31] http://www.peachpit.com/articles/article.aspx?p=1945331&seq-Num=1.

[32] http://www.readingsoft.com/.

[33] http://www.transparency.org/.

[34] http://www.whitehouse.gov/the_press_office/Transparencyan-dOpenGovernment.

[35] https://developer.apple.com/library/ios/documentation/UserExpe-rience/Conceptual/MobileHIG/.

[36] https://eagereyes.org/blog/2007/infovis-2007-infovis-for-the-mass-es.

[37] https://eagereyes.org/blog/2010/end-of-verifiable-com.

[38] https://eagereyes.org/blog/2010/trivialization-for-the-masses.

[39] https://eagereyes.org/criticism/the-rise-and-fall-of-swivel.

[40] https://okfn.org/opendata/.

[41] https://www.youtube.com/watch?v=aHxv_2BMJfw.

[42] On Mediapart: http://blogs.mediapart.fr/blog/la-redaction-de-me-

diapart/180314/co2-la-carte-de-la-pollution-mondiale; on visualizing.org: http://visualizing.org/visualizations/where-are-big-polluters-1971; on citylab.com: http://www.citylab.com/work/2014/03/map-historys-biggest-greenhouse-gas-polluters/8657/.

[43] http://appliedresearch.cancer.gov/areas/cognitive/item.html.

[44] http://bestwebgallery.com/category/html5/.

[45] http://codecanyon.net/item/pixel-map/full_screen_preview/1243869.

[46] http://collection.marijerooze.nl/#!/.

[47] http://d3js.org/.

[48] http://deputeviz.fr/other_apps/SI/wikipedia_template/example.php?t=1&s=2.

[49] http://deputeviz.fr/other_apps/SI/wikipedia_template/example.php?t=1&s=3.

[50] http://en.wikipedia.org/wiki/Help:Barchart.

[51] http://englishteststore.net/index.php?option=com_content&view=article&id=241&Itemid=285.

[52] http://genuineevaluation.com/the-risks-of-using-choropleth-maps/.

[53] http://hint.fm/projects/wind/.

[54] http://luna.cas.usf.edu/~mbrannic/files/pmet/irt.htm.

[55] http://nces.ed.gov/surveys/all/.

[56] http://opendefinition.org/od/.

[57] http://peopleviz.gforge.inria.fr/trunk/SI_examples/.

[58] http://perceptualedge.com/files/GraphDesignIQ.html.

[59] http://pushpoppress.com/ourchoice/.

[60] http://raphaeljs.com/.

[61] http://stuartroseman.com/post/10238550499/goodbye-to-verifiable-i-mean-it-this-time.

[62] http://tiny.cc/datastories.

[63] http://tinyurl.com/GoogleMaterial.

[64] http://tinyurl.com/GreenSmokeTut.

[65] http://tinyurl.com/NYTimes-collection.

[66] http://tinyurl.com/TrapcodeDoc.

[67] http://tinyurl.com/VisDataBestOfJan.

[68] http://tinyurl.com/WidnowsMetro.

[69] http://turkopticon.differenceengines.com/.

[70] http://visualizationliteracy.org/platform.

[71] http://worrydream.com/#!/MagicInk.

[72] http://www-969.ibm.com/software/analytics/manyeyes/#/visualizations/popular?view=201662.

[73] http://www-969.ibm.com/software/analytics/manyeyes/#/visualizations/popular?view=201665.

[74] http://www.ala.org/acrl/publications/whitepapers/presidential.

[75] http://www.awwwards.com/.

[76] http://www.babynamewizard.com/voyager#prefix=&sw=both&exact=false.

[77] http://www.brainpickings.org/2012/02/06/francesco-franchi-visual-storytelling/.

[78] http://www.cdc.gov/nchs/data/washington_group/meeting5/WG5_Appendix4.pdf.

[79] http://www.cdc.gov/physicalactivity/downloads/pa_state_indicator_report_2014.pdf.

[80] http://www.cesr.org/article.php?id=1756.

[81] http://www.csc.ncsu.edu/faculty/healey/PP/index.html#Tri_Psych_Review:88.

[82] http://www.deputeviz.fr/.

[83] http://www.donpotter.net/pdf/mwia.pdf.

[84] http://www.evoenergy.co.uk/uk-energy-guide/.

[85] http://www.evoenergy.co.uk/uk-energy-guide/.

[86] http://www.gapminder.org/.

[87] http://www.gapminder.org/downloads/.

[88] http://www.jnd.org/dn.mss/affordances_and.html.

[89] http://www.kent.ac.uk/careers/tests/mathstest.htm.

[90] http://www.medialit.org/reading-room/what-media-literacy-definitionand-more.

[91] http://www.michelbergerbooze.com/.

[92] http://www.nytimes.com/interactive/2012/11/02/us/politics/paths-to-the-white-house.html?_r=0.

[93] http://www.oecdbetterlifeindex.org/.

[94] http://www.quizrevolution.com/act101820/mini/go/.

[95] http://www.rasch-analysis.com/rasch-models.htm.

[96] http://www.regardscitoyens.org/.

[97] http://www.tableau.com/sites/default/files/media/which_chart_v6_final_0.pdf.

[98] http://www.theguardian.com/data.

[99] http://www.uis.unesco.org/Literacy/Pages/lamp-literacy-assessment.aspx.

[100] https://developers.google.com/chart/.

[101] https://github.com/INRIA/Visualization-Literacy-101.

[102] https://msdn.microsoft.com/en-us/library/aa733613(v=vs.60).aspx.

[103] https://okfn.org/opendata/.

[104] https://support.google.com/analytics/answer/1009409? hl=en.

[105] https://www.safaribooksonline.com/library/view/engaging-audiences-with/9781491909959/part00.html.

[106] https://www.youtube.com/playlist?list=PLAD53BD53C7A23E96.

[107] On Mediapart: http://blogs.mediapart.fr/blog/la-redaction-de-mediapart/180914/les-diplomes-est-ce-que-ca-paye; on visualizing.org: http://visualizing.org/visualizations/there-financial-interest-investing-longer-studies.

[108] On Mediapart: http://blogs.mediapart.fr/blog/la-redaction-de-mediapart/190914/le-poids-du-nucleaire-dans-le-monde; on visualizing.org: http://visualizing.org/visualizations/what-part-nuclear-energy.

[109] Debbie Abilock. Visual information literacy: Reading a documentary photograph. *Knowledge Quest*, 36(3), 2008.

[110] American Psychological Association. *The Publication manual of the American psychological association (6th ed.)*. Washington, DC, 2010.

[111] Simon Attfield, Gabriella Kazai, Mounia Lalmas, and Benjamin Piwowarski. Towards a science of user engagement. In *WSDM Workshop on User Modelling for Web Applications*. ACM International Conference on Web Search And Data Mining, 2011.

[112] Ronald Baecker, Ian Small, and Richard Mander. Bringing icons to life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '91, pages 1–6. Association for Computing Machinery (ACM), 1991.

[113] Justin Baer, Mark Kutner, and John Sabatini. *Basic reading skills and the literacy of the america's least literate adults: Results from the 2003 National Assess-*

ment of Adult Literacy (NAAL) supplemental studies. Technical report, National Center for Education Statistics (NCES), 2009.

[114] Albert Bandura. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review,* pages 191–215, 1977.

[115] Lyn Bartram, Colin Ware, and Tom Calvert. Moticons: detection, distraction and task. *International Journal of Human-Computer Studies,* 58(5):515–545, 2003.

[116] Scott Bateman, Regan Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. Useful junk? the effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '10, pages 2573–2582, 2010.

[117] Ruedi Baur. *Les 101 Mots du Design Graphique à l'Usage de Tous.* Archibooks, 2013.

[118] Jacques Bertin. *Semiologie Graphique: Les diagrammes, Les réseaux, Les cartes.* Éditions de l'École des Hautes Etudes en Sciences Sociales, 1999.

[119] Philip Bobko and Ronald Karren. The perception of Pearson product moment correlations from bivariate scatterplots. *Personnel Psychology,* 32(2):313–325, 1979.

[120] Rita Borgo, Alfie Abdul-Rahman, Farhan Mohamed, Philip W. Grant, Irene Reppa, Luciano Floridi, and Min Chen. An empirical study on using visual embellishments in visualization. *IEEE Transactions on Visualization and Computer Graphics,* 18 (12):2759–2768. Institute of Electrical and Electronics Engineers (IEEE), 2012.

[121] Michelle A. Borkin, Azalea A. Vo, Zoya Bylinskii,Phillip Isola,Shashank Sunkavalli, Aude Oliva, and Hanspeter Pfister. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics.*

Institute of Electrical and Electronics Engineers (IEEE), 2013.

[122] Jeremy Boy, Ronald A. Rensink, Enrico Bertini, and Jean-Daniel Fekete. A Principled Way of Assessing Visualization Literacy. *IEEE Transactions on Visualization and Computer Graphics.* Institute of Electrical and Electronics Engineers (IEEE), 2014.

[123] Jeremy Boy and Jean-Daniel Fekete. The CO2 Pollution Map: Lessons Learned from Designing a Visualization that Bridges the Gap between Visual Communication and Information Visualization. In *Proceedings of IEEE VIS Conference,* VIS '14. Institute of Electrical and Electronics Engineers (IEEE), 2014.

[124] Jeremy Boy, Louis Eveillard, Françoise Detienne, and Jean-Daniel Fekete. Suggested Interactivity: Seeking Perceived Affordances in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics.* Institute of Electrical and Electronics Engineers (IEEE), 2015.

[125] Jeremy Boy, Françoise Detienne, and Jean-Daniel Fekete. Storytelling in Information Visualizations: Does it Engage Users to Explore Data? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '15. Association for Computing Machinery (ACM), 2015.

[126] Valerie J. Bristor and Suzanne V. Drake. Linking the language arts and content areas through visual technology. *Technological Horizons in Education Journal,* 22(2):74- 77, 1994.

[127] William Buxton. A three-state model of graphical input. In *Proceedings of the IFIP TC13 Third Interational Conference on Human-Computer Interaction,* INTERACT '90, pages 449–456. North-Holland Publishing Co., 1990.

[128] Lee Byron and Martin Wattenberg. Stacked grgeometry & aesthetics. *IEEE Transactions on Visualization and Computer Graphics,* 14(6):1245–1252. Institute of Electrical and Electronics Engineers (IEEE), 2008.

[129] Tara C. Callaghan. Interference and Dominance in Texture Segregation: Hue, Geometric Form, and Line Orientation. *Perception and Psychopysics*, 46(4):299–311, 1989.

[130] Angelo Canty and Brian Ripley. *Boostrap Functions*, 2014.

[131] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman. *Readings in Information Visualization: Using Vision to Think.* Morgan Kaufmann Publishers Inc., 1999.

[132] Patricia Carpenter and Priti Shah. A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, 4(2): 75–100, 1998.

[133] Ronald P. Carver. *The Causes of Highand Low Reading Achievement.* Taylor & Francis, 2000.

[134] Hsiang Chen, Rolf T. Wigand, and Michael Sanford Nilan. Exploring web users' optimal flow experiences. *IT & People*, 13(4):263–281, 2000.

[135] Fanny Chevalier, Romain Vuillemot, and Guia Gali. Using Concrete Scales: A Practical Framework for Effective Visual Depiction of Complex Measures. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2426–2435. Institute of Electrical and Electronics Engineers (IEEE), 2013.

[136] Circulaire du 26 mai 2011 relative à la création du portail unique des informations publiques de l'État "data.gouv.fr" par la mission "Etalab" et l'application des dispositions régissant le droit de réutilisation des informations publiques. *Journal Officiel de la République Française*, 2011.

[137] William S. Cleveland, Persi Diaconis, and Robert McGill. Variables on scatterplots look more highly correlated when the scales are increased. *Science*, 216(4550): 1138–1141, 1982.

[138] William S. Cleveland and Robert McGill. Graphical perception: Theory,

experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):pp. 531–554, 1984.

[139] William S. Cleveland and Robert McGill. Graphical Perception and Graphical Methods for Analyzing Scientific Data. *Science*, 229:828–833, 1985.

[140] Michael Correll, Danielle Albers, Steven Franconeri, and Michael Gleicher. Comparing averages in time series data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1095–1104. ACM, 2012.

[141] Sir Geoffrey Crowther. The Crowther Report, volume 1. *Her Majesty's Stationery Office*, 1959.

[142] Mihály Csikszentmihalyi. *Beyond Boredom and Anxiety.* Jossey-Bass Publishers: The Jossey-Bass behavioral science series, 1975.

[143] Mihály Csikszentmihalyi and Rick E. Robinson. *The Art of Seeing: An Interpretation of the Aesthetic Encounter.* Getty Trust Publications: Getty Education Institute for the Arts Series. J.P. Getty Museum, 1990.

[144] Mihály Csikszentmihalyi. *Flow: The Psychology of Optimal Experience.* Harper Perennial, 1991.

[145] Geoff Cumming. The new statistics: Why and how. *Psychological science*, 25(1):7–29, 2014.

[146] Frances R. Curcio. Comprehension of mathematical relationships expressed in graphs. *Journal for research in mathematics education*, pages 382–393, 1987.

[147] Michael J. Danziger. *Information Visualization for the People.* Master's thesis, Massachusetts Institute of Technology, Department of Comparative Media Studies, 2008.

[148] Edward Deci and Richard M. Ryan. *Perspectives in Social Psychology.* Springer

US, 1985.

[149] Nicholas Diakopoulos. Game-y information graphics. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems,* CHI EA '10, pages 3595–3600. Association for Computing Machinery (ACM), 2010.

[150] Nicholas Diakopoulos, Funda Kivran-Swaine, and Mor Naaman. Playable data: Characterizing the design space of game-y infographics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '11, pages 1717– 1726. Association for Computing Machinery (ACM), 2011.

[151] Daniel Dietrich, Jonathan Gray, Tim McNamara, Antti Poikola, Rufus Pollock, Julian Tait, and Ton Zijlstra. *The Open Data Handbook.* 2012.

[152] Daniel Dorling. Area Cartograms: Their Use and Creation. *Concepts & Techniques in Modern Geography.* Environmental Publications, 1996.

[153] Pierre Dragicevic, Fanny Chevalier, and Stéphane Huot. Running an HCI Experiment in Multiple Parallel Universes. *In CHI '14 Extended Abstracts on Human Factors in Computing Systems,* CHI EA '14, pages 607–618. Association for Computing Machinery (ACM), 2014.

[154] Ronald S. Easterby. The perception of symbols for machine displays. *Ergonomics,* 13(1):149–158, 1970. PMID: 5416867.

[155] Jean-Daniel Fekete, Jarke J.van Wijk, John T. Stasko, and Chris North. The Value of Information Visualization. *Information visualization,* pages 1–18. Springer Verlag, 2008.

[156] Stephen Few. Infovis as seen by the world out there: 2007 inreview. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology,* VAST '07. Institute of Electrical and Electronics Engineers (IEEE), 2007.

[157] Stephen Few. Benefitting InfoVis with Visual Difficulties? Provocation Without a Cause, 2011.

[158] Stephen Few. The Chartjunk Debate–A Close Examination of Recent Findings, 2011.

[159] R. Chris Fraley, Niels G. Waller, and Kelly A. Brennan. An item response theory analysis of self-report measures of adult attachment. *Journal of personality and social psychology*, 78(2):350, 2000.

[160] Eric G. Freedman and Priti Shah. Toward a model of knowledge-based graph comprehension. In Mary Hegarty, Bernd Meyer, and N. Hari Narayanan, editors, *Diagrammatic Representation and Inference*, volume 2317 of Lecture Notes in Computer Science, pages 18–30. Springer Berlin Heidelberg, 2002.

[161] Jill Freyne and Barry Smyth. Visualization for the masses: Learning from the experts. In Isabelle Bichindaritz and Stefania Montani, editors, Case-Based Reasoning. *Research and Development*, volume 6176 of Lecture Notes in Computer Science, pages 111–125. Springer Berlin Heidelberg, 2010.

[162] Susan N. Friel, Frances R. Curcio, and George W. Bright. Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in mathematics Education*, 32(2):124–158, 2001.

[163] James J. Gibson. *The Ecological Approach to Visual Perception*. Resources for ecological psychology. Erlbaum, 1986.

[164] Stephen Gilroy, Julie Porteous, Fred Charles, and Marc Cavazza. Exploring passive user interaction for adaptive narratives. In *Proceedings of the ACM International Conference on Intelligent User Interfaces*, IUI '12, pages 119–128. Association for Computing Machinery (ACM), 2012.

[165] David Gittins. Icon-based human-computer interaction. *International Journal of Man-Machine Studies*, 24(6):519–543, 1986.

[166] David Gotz and Zhen Wen. Behavior-driven visualization recommenda-

tion. In *Proceedings of the ACM International Conference on Intelligent User Interfaces,* IUI ’09, pages 315–324. Association for Computing Machinery (ACM), 2009.

[167] Arthur C. Graesser, Shane S. Swamer, William B. Baggett, and Marie A. Sell. New models of deep comprehension. *Models of understanding text,* pages 1–32, 1996.

[168] Chris Harrison, Gary Hsieh, Karl D.D. Willis, Jodi Forlizzi, and Scott E. Hudson. Kineticons: using iconographic motion in graphical user interface design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI ’11, pages 1999–2008. Association for Computing Machinery (ACM), 2011.

[169] H. Rex Hartson. Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour and Information Technology,* 22:315–338, 2003.

[170] Christopher G. Healey. *Visualization of Multivariate Data Using Preattentive Processing.* Master’s thesis, University of British Columbia, 1992.

[171] Christopher G. Healey, Kellogg S. Booth, and James T. Enns. Visualizing real-time multivariate data using preattentive processing. *ACM Transactions on Modeling and Computer Simulation,* 5(3):190–221. Association for Computing Machinery (ACM), 1995.

[172] Jeffrey Heer. Socializing visualization. In *CHI ’06 Extended Abstracts on Human Factors in Computing Systems,* CHI EA ’06. Association for Computing Machinery (ACM), 2006.

[173] Jeffrey Heer, Fernanda Viégas, and Martin Wattenberg. Voyagers and voyeurs: Supporting asynchronous collaborative information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI ’07, pages 1029–1038, 2007.

[174] Jeffrey Heer. Visualization evaluation of the masses, by the masses, and for the masses. In *Proceedings of the 3rd BELIV'10 Workshop: BEyond time and errors: novel evaLuation methods for Information Visualization*, BELIV '10, 2010.

[175] Jeffrey Heer and Mike Bostock. Crowdsourcing graphical perception: using Mechanical Turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '10, pages 203–212. Association for Computing Machinery (ACM), 2010.

[176] Jeffrey Heer and Ben Shneiderman. Interactive dynamics for visual analysis. *Communications ACM,* 55(4):45–54. Association for Computing Machinery (ACM), 2012.

[177] Shih-Miao Huang, Kong-King Shieh, and Chai-Fen Chi. Factors affecting the design of computer icons. *International Journal of Industrial Ergonomics,* pages 211– 218, 2002.

[178] Daniel E. Huber and Christopher G. Healey. Visualizing data with motion. In *Proceedings of the VisWeek Conference,* VisWeek '05, page 67. Institute of Electrical and Electronics Engineers (IEEE), 2005.

[179] Jessica Hullman, Eytan Adar, and PritiShah. Benefiting infovis with visual difficulties. *IEEE Transactions on Visualization and Computer Graphics,* 17(12):2213–2222. Institute of Electrical and Electronics Engineers (IEEE), 2011.

[180] Samuel Huron, Yvonne Jansen, and Sheelagh Carpendale. Constructing Visual Representations: Investigating the Use of Tangible Tokens. *IEEE Transactions on Visualization and Computer Graphics.* Institute of Electrical and Electronics Engineers (IEEE), 2014.

[181] Petra Isenberg, Anastasia Bezerianos, Pierre Dragicevic, and Jean-Daniel Fekete. A study on dual-scale data charts. *IEEE Transactions on Visualization and Computer Graphics,* 17(12):2469–2478. Institute of Electrical and

Electronics Engineers (IEEE), 2011.

[182] Manfred Jahn. *Narratology: A Guide to the Theory of Narrative.* University of Cologne, 2005.

[183] Yvonne Jansen and Pierre Dragicevic. An Interaction Model for Visualizations Beyond The Desktop. *IEEE Transactions on Visualization and Computer Graphics,* 19 (12):2396–2405. Institute of Electrical and Electronics Engineers (IEEE), 2013.

[184] Morgan Jennings. Theory and models for creating engaging and immersive e-commerce websites. In *Proceedings of the SIGCPR Conference,* SIGCPR '00, pages 77–85. Association for Computing Machinery (ACM), 2000.

[185] Ross Keith. Growing Open Data Movement Could Mean Big Bucks. 2014.

[186] Irwin Kirsch. *The International Adult Literacy Survey (IALS): Understanding what was measured.* Technical report, Educational Testing Service, 2001.

[187] Nichole Knutson, Kathryn Shirley Akers, and Kelly D Bradley. Applying the Rasch model to measure first-year students' perceptions of college academic readiness. In *Proceedings of the Annual Meeting of the MWERA*, 2010.

[188] Scott D. Kominers. Sticky content and the structure of the commercial web. In *Proceedings of the NetEcon Workshop on the Economics of Networks, Systems and Computation*, NetEcon'09, 2009.

[189] Mitchel Langford and David J. Unwin. Generating and mapping population density surfaces within a geographical information system. *The Cartographic Journal,* 31(1):21–26, 1994.

[190] Andrea Lau and Andrew Vande Moere. Towards a model of information aesthetics in information visualization. In *Proceedings of the 11th International Conference Information Visualization,* IV '07, pages 87–92, 2007.

[191] Adam Light and Patrick J. Bartlein. The end of the rainbow? color schemes for improved data graphics. *Eos, Transactions American Geophysical Union,* 85(40): 385–391, 2004.

[192] Chao Liu, Ryen W. White, and Susan Dumais. Understanding web browsing behaviors through weibull analysis of dwell time. *In Proceedings of the 33rd International SIGIR Conference on Research and Development in Information Retrieval,* SIGIR '10, pages 379–386. Association for Computing Machinery (ACM), 2010.

[193] Sara Ljungblad and Tobias Skog. Integrating information with aesthetics: Ambient infovis. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems,* CHI EA '04. Association for Computing Machinery (ACM), 2004.

[194] Matt Lockyer, Lyn Bartram, and Bernhard E. Riecke. Simple motion textures for ambient affect. In *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging,* CAe '11, pages 89–96. Association for Computing Machinery (ACM), 2011.

[195] Kenneth N. Lodding. Iconic interfacing. *IEEE Computer Graphics and Applications,* 3(2):11–20. Institute of Electrical and Electronics Engineers (IEEE), 1983.

[196] Richard Lowe. "Reading" scientific diagrams: Characterising components of skilled performance. *Research in Science Education,* 18(1):112–122, 1988.

[197] Richard K. Lowe. *Scientific diagrams: How well can students read them? what research says to the science and mathematics teacher.* Report, microform, online, Curtin University of Technology, Perth (Australia). National Key Centre for Science and Mathematics., 1989.

[198] Kendra Mack. *Viral visualizations: Online infographics as reflections of the internet's informational landscape.* 2011.

[199] Jock Mackinlay. *Automating the design of graphical presentations of relational information.* ACM Transactions on Graphics, 5(2):110–141. Association for Computing Machinery (ACM), 1986.

[200] Stephen Manes. Pushing picture-perfect programs: Smash that icon! *PC Magazine,* page 64, 1985.

[201] Jennifer Mankoff, Anind K. Dey, Gary Hsieh, Julie Kientz, Scott Lederer, and Morgan Ames. Heuristic evaluation of ambient displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '03, pages 169–176. Association for Computing Machinery (ACM), 2003.

[202] Gary Marchionini. *Information Seeking in Electronic Environments.* Cambridge Series on Human-Computer Interaction. Cambridge University Press, 1997.

[203] Étienne-Jules Marey. *La méthode graphique dans les sciences expérimentales et principalement en physiologie et en médecine.* G. Masson, 1878.

[204] John McCarthy and Peter Wright. Technology as experience. *Interactions,* 11(5): 42–43, 2004.

[205] Joachim Meyer, Meirav Taieb, and Ittai Flascher. Correlation estimates as perceptual judgments. *Journal of Experimental Psychology: Applied,* 3(1):3, 1997.

[206] Steve Mollman. When it comes to making data sexy, you can't be too graphic. 2009.

[207] Tamara Munzner. *Visualization Analysis and Design.* A. K. Peters Visualization Series. Taylor & Francis, 2014.

[208] Peter Neri, M. Concetta Morrone, and David C. Burr. Seeing biological motion. *Nature,* 395(6705):894–896, 1998.

[209] Moritz Neumüller. *Hypertext Semiotics in the Commercialized Internet.* PhD thesis, 2001.

[210] Otto Neurath. *Visual Education: A New Language.* 1937.

[211] Donald A. Norman. *The design of every day things.* Basic Books, 2002.

[212] Chris North. Toward measuring visualization insight. *IEEE Computer Graphics and Applications,* 26(3):6–9. Institute of Electrical and Electronics Engineers (IEEE), 2006.

[213] Heather L. O'Brien and Elaine G. Toms. What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology* 59(6):938–955, 2008.

[214] Heather L. O'Brien and Elaine G. Toms. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology* 61(1):50–69, 2010.

[215] Heather L O'Brien. Exploring user engagement in online news interactions. *Journal of the American Society for Information Science and Technology* 48(1):1– 10, 2011.

[216] OECD. *Pisa 2012 assessment and analytical framework.* Technical report, OECD, 2012.

[217] OTA. Computerized Manufacturing Automation: Employment, Education, and the Workplace. *United States Office of technology Assessment,* 1984.

[218] Lori McCay-Peet, Mounia Lalmas, and Vidhya Navalpakkam. On saliency, affect and focused attention. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '12, pages 541–550. Association for Computing Machinery (ACM), 2012.

[219] Charles Perin, Romain Vuillemot, and Jean-Daniel Fekete. À table!: Improving temporal navigation in soccer ranking tables. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '14, pages 887–896. Association for Computing Machinery (ACM), 2014.

[220] Charles Perin, Pierre Dragicevic, and Jean-Daniel Fekete. Revisiting Bertin matrices: New Interactions for Crafting Tabular Visualizations. *IEEE Transactions on Visualization and Computer Graphics.* Institute of Electrical and Electronics Engineers (IEEE), 2014.

[221] Steven Pinker. *A Theory of Graph Comprehension.* Center for Cognitive Science Cambridge, Massachusetts: Occasional paper, 1981.

[222] Peter Pirolli and Stuart K. Card. *The evolutionary ecology of information foraging.* Technical report, 1997.

[223] William Playfair. *Playfair's Commercial and Political Atlas and Statistical Breviary.* Cambridge University Press, 2005.

[224] Irwin Pollack. Identification of visual correlational scatterplots. *Journal of Experimental Psychology,* 59(6):351, 1960.

[225] Zachary Pousman and John T. Stasko. A taxonomy of ambient information systems: Four patterns of design. In *Proceedings of the Working Conference on Advanced Visual Interfaces,* AVI '06, pages 67–74. Association for Computing Machinery (ACM), 2006.

[226] Zachary Pousman, John T. Stasko, and Michael Mateas. Casual information visualization: Depictions of data in everyday life. *IEEE Transactions on Visualization and Computer Graphics,* 13(6):1145–1152. Institute of Electrical and Electronics Engineers (IEEE), 2007.

[227] Zenon Pylyshyn, Jacquelyn Burkell, Brian Fisher, Christopher Sears, William Schmidt, and Lana Trick. Multiple parallel access in visual attention. *Canadian Journal of Experimental Psychology,* 48(2):260–83, 1994.

[228] Raj M. Ratwani and J. Gregory Trafton. Shedding light on the graph schema: Perceptual features versus invariant structure. *Psychonomic Bulletin and Review,* 15 (4):757–762, 2008.

[229] William T. Reeves. Particle systems—A technique for modeling a class of fuzzy objects. *ACM Transactions on Graphics,* 2(2):91–108. Association for Computing Machinery (ACM), 1983.

[230] Ronald A. Rensink and Gideon Baldridge. The perception of correlation in scatterplots. *Computer Graphics Forum,* 29(3):1203–1210, 2010.

[231] François Richaudeau. *Méthode de Lecture rapide.* Retz, 2004.

[232] Dimitris Rizopoulos. ltm: An R package for latent variable modeling and Item Response Analysis. *Journal of Statistical Software,* 17(5):1–25, 2006.

[233] Kerry Rodden, Hilary Hutchinson, and Xin Fu. Measuring the user experience on a large scale: User-centered metrics for web applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '10, 2010.

[234] Joel Ross, Lilly Irani, M. Six Silberman, Andrew Zaldivar, and Bill Tomlinson. Who are the crowdworkers? Shifting demographics in mechanical turk. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems,* CHI EA '10, pages 2863– 2872. Association for Computing Machinery (ACM), 2010.

[235] Richard M. Ryan and Edward L. Deci. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology,* 25(1):54–67, 2000.

[236] Arvind Satyanarayan and Jeffrey Heer. Authoring narrative visualizations with ellipsis. In *Proceedings of the Eurographics/IEEE VGTC Conference on Visualization,* EuroVis '14. Institute of Electrical and Electronics Engineers (IEEE), 2014.

[237] Brian J. Scholl and Patrice D. Tremoulet. Perceptual causality and animacy. *Trends Cognitive Sciences,* 4(8):299–309, 2000.

[238] Edward Segel and Jeffrey Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics,* 16(6):1139–1148. Institute of Electrical and Electronics Engineers (IEEE), 2010.

[239] Priti Shah. A model of the cognitive and perceptual processes in graphical display comprehension. *Reasoning with diagrammatic representations,* pages 94–101, 1997.

[240] Ben Shneiderman. Direct manipulation: A step beyond programming languages. *IEEE Computer,* 16(8):57–69. Institute of Electrical and Electronics Engineers (IEEE), 1983.

[241] Ben Shneiderman. Tree visualization with tree-maps: 2-d space-filling approach. *ACM Transactions on Graphics,* 11(1):92–99. Association for Computing Machinery (ACM), 1992.

[242] Ben Shneiderman. Dynamic queries for visual information seeking. *IEEE Software,* 11(6):70–77. Institute of Electrical and Electronics Engineers (IEEE), 1994.

[243] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages,* VL '96, page 336. Institute of Electrical and Electronics Engineers (IEEE), 1996.

[244] Tobias Skog, Sara Ljungblad, and Lars Erik Holmquist. Between aesthetics and utility: Designing ambient information visualizations. In *Proceedings of the VisWeek Conference,* VisWeek '03, pages 233–240. Institute of Electrical and Electronics Engineers (IEEE), 2003.

[245] Robert Spence. *Information Visualization: Design for Interaction (2nd ed.).* Prentice-Hall, Inc., 2007.

[246] John T. Stasko, Todd Miller, Zachary Pousman, Christopher Plaue, and Osman Ullah. Personalized peripheral information awareness through information art. In Nigel Davies, ElizabethD. Mynatt, and Itiro Siio, editors, *UbiComp 2004: Ubiquitous Computing,* volume 3205 of Lecture Notes in Computer Science, pages 18–35. Springer Berlin Heidelberg, 2004.

[247] Jian Tang, Yuxiang Zhao, and Ping Zhang. Perceived affordances of web advertisements: Implications for information artifacts design. In *Proceedings of the Fifth China Summer Workshop on Information Management,* CSWIM, 2011.

[248] Laura G. Tateosian, Christopher G. Healey, and James T. Enns. Engaging viewers through nonphotorealistic visualizations. In *Proceedings of the 5th International Symposium on Non-Photorealistic Animation and Rendering,* NPAR '07, pages 93– 102. Association for Computing Machinery (ACM), 2007.

[249] Conrad Taylor. New kinds of literacy, and the world of visual information. *Literacy,* 2003.

[250] Elaine G. Toms. Information interaction: Providing a framework for information architecture. *Journal of the American Society for Information Science and Technology,* 53(10):855–862, 2002.

[251] Noam Tractinsky, Adi S. Katz, and Dror Ikar. What is beautiful is usable. *Interacting with Computers,* 13(2):127–145, 2000.

[252] J. Gregory Trafton, Sandra P. Marshall, Farilee Mintz, and Susan B. Trickett. Extracting explicit and implict information from complex visualizations. In Mary Hegarty, Bernd Meyer, and N. Hari Narayanan, editors, *Diagrammatic Representation and Inference,* volume 2317 of Lecture Notes in Computer Science, pages 206– 220. Springer Berlin Heidelberg, 2002.

[253] Susan B. Trickett and J. Gregory Trafton. Toward a comprehensive model of graph comprehension: Making the case for spatial cognition. In Dave

Barker-Plummer, Richard Cox, and Nik Swoboda, editors, *Diagrammatic Representation and Inference,* volume 4045 of Lecture Notes in Computer Science, pages 286–300. Springer Berlin Heidelberg, 2006.

[254] Jan Tschichold. *Livre et Typographie.* Allia, 1998.

[255] Edward R. Tufte. *The Visual Display of Quantitative Information.* Graphics Press, 1986.

[256] Barbara Tversky. What do sketches say about thinking? In T. Stahovic, J. Landay, and R. Davis, editors, *AAAI Spring Symposium on Sketch Understanding.* AAAI Press, 2002.

[257] Barbara Tversky, Julie Bauer Morrison Y., and Mireille Betrancourt. Animation: Can it facilitate. *International Journal of Human-Computer Studies,* 57:247–262, 2002.

[258] Barbara Tversky, Julie Bauer Morrison, and Mireille Betrancourt. Animation: Can it facilitate. *International Journal of Human-Computer Studies,* 57:247–262, 2002.

[259] Barbara Tversky. Semantics, Syntax, and Pragmatics of graphics. *Language and Visualization,* pages 141–158. Lund University Press, 2004.

[260] Barbara Tversky. *Visuospatial reasoning,* chapter 10, pages 209–249. Cambridge University Press, 2005.

[261] Barbara Tversky. *Communicating with diagrams and gestures.* Macmillan, 2007.

[262] Sandra Utt and Steve Pasternack. Infographics today: Using qualitative devices to display quantitative information. *Newspaper research journal,* 14(3):146–157, 1993.

[263] Jarke J. van Wijk. The value of visualization. In *Proceedings of the VisWeek*

Conference, VisWeek '05, page 11. Institute of Electrical and Electronics Engineers (IEEE), 2005.

[264] Jo Vermeulen, Kris Luyten, Elise van den Hoven, and Karin Coninx. Crossing the bridge over norman's gulf of execution: Revealing feedforward's true identity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, 2013. Association for Computing Machinery (ACM), 2012.

[265] Fernanda Viégas and Martin Wattenberg. Communication-minded visualization: A call to action [technical forum]. *IBM Systems Journal*, 45(4):801–812, 2006.

[266] Fernanda Viégas and Martin Wattenberg. Artistic datavisualization: Beyond visual analytics. In Douglas Schuler, editor, *HCI (15)*, volume 4564 of Lecture Notes in Computer Science, pages 182–191. Springer, 2007.

[267] Fernanda Viégas, Martin Wattenberg, Frank van Ham, Jesse Kriss, and Matt McKeon. Manyeyes: a site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1121–1128. Institute of Electrical and Electronics Engineers (IEEE), 2007.

[268] Howard Wainer. A test of graphicacy in children. *Applied Psychological Measurement*, 4(3):331–340, 1980.

[269] Martin Wattenberg. Baby names, visualization, and social data analysis. In *Proceedings of the VisWeek Conference*, VisWeek '05, pages 1–7. Institute of Electrical and Electronics Engineers (IEEE), 2005.

[270] Martin Wattenberg and Jesse Kriss. Designing for social data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):549–557. Institute of Electrical and Electronics Engineers (IEEE), 2006.

[271] Wesley Willett, Jeffrey Heer, Joseph Hellerstein, and Maneesh Agrawala. Commentspace: Structured support for collaborative visual analysis. In

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11. Association for Computing Machinery (ACM), 2011.

[272] Wesley Willett, Jeffrey Heer, and Maneesh Agrawala. Strategies for crowdsourcing social data analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* CHI '12, pages 227–236. Association for Computing Machinery (ACM), 2012.

[273] Wesley Willett, Shiry Ginosar, Avital Steinitz, Björn Hartmann, and Maneesh Agrawala. Identifying Redundancy and Exposing Provenance in Crowdsourced Data Analysis. *IEEE Transactions on Visualization and Computer Graphics,* 19(12): 2198–2206. Institute of Electrical and Electronics Engineers (IEEE), 2013.

[274] Wendy L. Winn. *Visualizing Science: A Semiotic Analysis of Visual Representations in Ornithology Journals, 1859–2003.* PhD thesis, University of Minnesota, 2006.

[275] Jo Wood, Petra Isenberg, Tobias Isenberg, Jason Dykes, Nadia Boukhelifa, and Aidan Slingsby. Sketchy rendering for information visualization. *IEEE Transactions on Visualization and Computer Graphics,* 18(12):2749–2758. Institute of Electrical and Electronics Engineers (IEEE), 2012.

[276] John K. Wright. A method of mapping densities of population: with Cape Cod as an example. *Geographical Review,* 26(1):pp. 103–110, 1936.

[277] Margaret Wu, and Ray Adams. Applying the Rasch model to psycho-social measurement [electronic resource]: A practical approach/Margaret Wu and Ray Adams. *Educational Measurement Solutions,* Melbourne, Vic, 2007.

[278] Ji Soo Yi, Youn ah Kang, John T. Stasko, and Julie A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics,* 13(6):1224–1231. Institute of Electrical and Electronics Engineers (IEEE), 2007.

Engager les Citoyens à Aller au-delà des Simples Représentations de Données Ouvertes

Jeremy BOY

RESUME : Dans ce manuscrit, j'étudie quatre obstacles potentiels à l'engagement d'un internaute avec une interface de visualisation d'informations interactive. Ma question de recherche principale est : comment ces obstacles sont-ils susceptibles de limiter l'engagement de l'utilisateur dans l'exploration efficace de données et comment remédier à ces limitations ? Je définis les quatre obstacles en termes de sous-coûts de la perception et de l'exploration en me référant au modèle proposé par van Wijk ; ils sont : 1) un coût de littératie, 2) un coût d'interprétation du contexte, 3) un coût de perception d'interactivité et 4) un coût de motivation initiale à explorer des données. Pour chacun, j'adopte soit une approche expérimentale pour mesurer le coût en question, soit une approche design pour aider les internautes à le surmonter. J'évalue aussi l'effet de certains éléments de design de visualisation reconnus pour leurs qualités communicationnelles sur l'engagement des internautes à explorer des données.

MOTS-CLEFS : Visualisation de données, Infographie, Visualisation d'informations, Visualisation pour les masses, Visualisation narrative, Littératie en visualisation, Affordances perçues, Suggestion de l'interactivité, Communication visuelle, Données ouvertes, Open data.

ABSTRACT : In this dissertation, I explore four initial challenges an online user may encounter when confronted with an information visualization website. The main research question I address is : how might these challenges limit people's engagement in efficient explorations of data, and how might these limitations be remedied ? I define the four challenges in terms of sub-costs of van Wijk's perception and exploration costs ; these are : 1) a literacy cost, 2) a context-interpretation cost, 3) a perceived interactivity cost, and 4) an initial incentive for exploration cost. For each, I propose either a way to assess or a method to help overcome the sub-cost. I also investigate whether popular techniques recommended for making visualizations engaging outside of purely analytical contexts can lead online users to explore data.

KEY-WORDS : Data visualisation, Infographics, Information visualization, Information visualization for the People, Narrative visualization, Visualization literacy, Perceived affordances, Suggested interactivity, Visual communication, Open data.